

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comp-gen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2003, 16:33:04 : Search time 735.169 Seconds  
(without alignments)  
10633.850 Million cell updates/sec

Title: US-09-856-681-1

Perfect score: 3093  
Sequence: 1 atgagtcagaagccttgat.....ccaatgctgctgtacataa 3093

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Capext 1.0

Searched: 2552756 seqs, 13497:9017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03:\*

```

1: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDSI/qcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY:HS

Result No.	Score	Query Match	Length	DB ID	Description
1	3093	100.0	3862	21	AA001233 Human semaphorin 6
2	3038	98.2	3498	21	AAA93617 Human semaphorin p
3	3032	98.0	4280	25	ABX71103 Novel human cDNA s
4	2969.6	96.0	6060	22	AAH18729 Human cDNA sequenc
5	2746.2	88.8	3550	19	AAV44295 Human secreted pro
6	2746.2	88.8	3550	22	AAF598469 Human cDNA clone c
7	2698	87.2	3333	21	AAA93618 Human semaphorin p
8	2221.8	71.8	3039	23	AA568807 DNA encoding novel

	9	2221.8	71.8	3039	23	AA568807	DNA encoding novel
10	2182.4	70.6	3041	22	AAH17625	Human cDNA sequenc	
11	1881.6	60.8	1890	21	AA93630	Human semaphorin p	
12	1824.4	59.0	2293	22	AAH15834	Human cDNA sequenc	
13	1742.4	56.3	2227	22	AAK94365	Human full-length	
14	1656.4	53.9	2306	22	AAH17861	Human cDNA sequenc	
15	1592	51.5	2592	23	AA568253	DNA encoding novel	
16	1462	47.3	2123	22	AAH13955	Human cDNA sequenc	
17	1216	39.3	2262	24	ABK34739	Human cDNA for nov	
18	962	31.1	1054	20	AAK84066	Human ZSMF-3 codin	
19	808.4	26.1	1047	20	AAK84076	Mouse ZSMF-3 codin	
20	788	25.5	839	22	AAH05402	Human cDNA clone (	
21	758.6	24.5	814	22	AAK92527	Human cDNA 5'-end	
22	758.6	24.5	814	22	AAK93383	Human cDNA clone r	
23	751.2	24.3	780	22	AAH05580	Human cDNA clone (	
24	742.6	24.0	788	22	AAK93916	Human cDNA clone r	
25	708.8	22.9	963	20	AAK84067	Degenerate Human Z	
26	686.8	22.2	963	20	AAK84077	Degenerate Mouse Z	
27	669	21.6	4234	24	ABAC0055	CADHP-2 coding seq	
28	651.6	21.4	3205	24	AB564380	Human semaphorin 6	
29	657.4	21.3	662	22	AAH08370	Human cDNA clone (	
30	657.4	21.3	1923	22	AAH42598	Partial cDNA sequenc	
31	657.4	21.3	3694	22	AAH42597	Nucleotide sequenc	
32	642	20.8	2191	24	AB564381	Human semaphorin-i	
33	642	20.8	2359	24	AB564383	Human semaphorin-1	
34	642	20.8	3196	24	AB564382	Human semaphorin-1	
35	642	20.8	3364	24	AB564384	Human semaphorin-1	
36	640.6	20.7	889	23	AA568806	DNA encoding novel	
37	577.4	18.7	1896	24	AB564379	Human semaphorin-1	
38	577.4	18.7	2014	24	AAH38696	Human LP221 secret	
39	570.8	18.5	786	22	AAH05233	Human cDNA clone (	
40	566.4	18.3	6645	25	ABX34714	Human mdmt cDNA SE	
41	565	18.3	1472	21	AAK98050	Human colon cancer	
42	536.6	18.0	2506	25	AB235865	Human secretory po	
43	535.4	18.0	3524	18	AAH72108	Human semaphorin 2	
44	535.4	18.0	3524	25	AAK08822	Angiogenesis-assoc	
45	533.8	17.9	3721	21	AAA96343	cDNA encoding a no	

# ALIGNMENTS

RESULT 1

AA001233  
ID AA001233 standard; DNA: 3862 BP.

AC AA001233;

DT 04-OCT-2000 (first entry)

DE Human semaphorin 6A-1 cDNA.

KW Human semaphorin 6A-1; (HSA)SEMA6A-1; neuronal development; apoptosis;  
neuronal regeneration; Fna/VASP protein family; immunomodulatory;  
KW gene therapy; diagnostic agent; therapeutic agent; differentiation;  
KW cytoskeletal stabilisation; plasticity; ds.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 658...3750

FT /\*tag= a

FT /\*product= "Human semaphorin 6A-1"

FT /\*note= "This region is specifically claimed as

FT SEQ ID NO: 1 in claim 1"

FT misc\_feature 3532...3747

FT /\*tag= b

FT /\*note= "Encodes the binding domain of semaphorin

FT 6A-1 which selectively binds to members of Fna/VASP

FT protein family"

XX WO200031252-A1.



Db 2278 GTTAAAGAACCTGATTGCTCCAGAGACCAATATTGGATGCATAAAGAGGTTGGT 2277  
QY 1621 GCGTCAGGCATTATCACCCCAACAGCAGACTGACTTTTGGCAGGACATAGAGCGTGGC 1680  
Db 2278 GCGTCAGGCATTATCACCCCAACAGCAGACTGACTTTTGGCAGGACATAGAGCGTGGC 2337  
QY 1681 ATACAGATGGGTCTGGGGGAGCTGTACAAATTCCTTTTGGCAGCTGAATGGGCATTCAGT 1740  
Db 2338 AATACAGATGGGTCTGGGGGAGCTGTACAAATTCCTTTTGGCAGCTGAATGGGCATTCAGT 2397  
QY 1741 TCCCTTCCCGCAGCAGACACACATACAGATTCGACGGCTCAAGAGGGGTATGAGTCTAGG 1806  
Db 2398 TCCCTTCCCGCAGCAGACACACATACAGATTCGACGGCTCAAGAGGGGTATGAGTCTAGG 2457  
QY 1801 GGAGGAATGCTGGAGTGGAGGATCTGTTGACTCACTGACAGACAGACCCGTTGGGG 1860  
Db 2458 GGAGGAATGCTGGAGTGGAGGATCTGTTGACTCACTGACAGACAGACCCGTTGGGG 2517  
QY 1861 GCAGTGTCTTCCCAATAATCACTCAAGACAGAGAGGAGTGAATCGGGAAGTACTGAAA 1920  
Db 2518 GCAGTGTCTTCCCAATAATCACTCAAGACAGAGAGGAGTGAATCGGGAAGTACTGAAA 2577  
QY 1921 GGCACGACACAGTGTCTCCCTGACCCCTCTGGCCATTCGACAGTCACTGCTGCTTCGTC 1980  
Db 2578 GGCACGACACAGTGTCTCCCTGACCCCTCTGGCCATTCGACAGTCACTGCTGCTTCGTC 2537  
QY 1981 ATGGGGCGCGTCTCTGGGGATACCGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040  
Db 2638 ATGGGGCGCGTCTCTGGGGATACCGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2697  
QY 2041 GTGGCTGTGTGTCAGCGCAAGAGAGAGCTCACCCACTCGCGCGGGGTCTCATGAGC 2100  
Db 2698 GTGGCTGTGTGTCAGCGCAAGAGAGAGCTCACCCACTCGCGCGGGGTCTCATGAGC 2757  
QY 2101 AGCGTCACCAAGTCTAGCGGCTCTTTTGGGACATCTAATCCAAAGACCCAAAGCGCGAG 2160  
Db 2758 AGCGTCACCAAGTCTAGCGGCTCTTTTGGGACATCTAATCCAAAGACCCAAAGCGCGAG 2817  
QY 2161 GCATCTCTCAGCCACTCATGCACACAGCAGCTCGCCACTCCCGGCAACACAGCCCAAG 2220  
Db 2818 GCATCTCTCAGCCACTCATGCACACAGCAGCTCGCCACTCCCGGCAACACAGCCCAAG 2877  
QY 2221 ATGCTCATTAAGCAGACACACACCTGGACCTGAGCGGCTCCCGACCCCGCAGAGTCA 2280  
Db 2878 ATGCTCATTAAGCAGACACACACCTGGACCTGAGCGGCTCCCGACCCCGCAGAGTCA 2937  
QY 2281 ACCCAACGCTGACGNGAATGGAGAGCCACCGGACAGCGGAGTGGAGAGAGAC 2340  
Db 2938 ACCCAACGCTGACGNGAATGGAGAGCCACCGGACAGCGGAGTGGAGAGAGAC 2997  
QY 2341 CAGAACCTCATCAATGCTGTACAAAGAGGACATGGCCGCTGCTGCTGCTGCTGCTGCTGCT 2400  
Db 2998 CAGAACCTCATCAATGCTGTACAAAGAGGACATGGCCGCTGCTGCTGCTGCTGCTGCTGCT 3057  
QY 2401 ACCGAGCTGCGGCTGCGGCGTCCCGGACCCACATCCCGAGCGTGGTGTCTGCGCCATC 2460  
Db 3058 ACCGAGCTGCGGCTGCGGCGTCCCGGACCCACATCCCGAGCGTGGTGTCTGCGCCATC 3117  
QY 2461 ACCGAGCGGCTACCGATGAGTACGTGGACGCGCAAAATGAGCGAGTGGCCAG 2520  
Db 3118 ACCGAGCGGCTACCGATGAGTACGTGGACGCGCAAAATGAGCGAGTGGCCAG 3177  
QY 2521 ATGGCGCTGGAGGACCGCGCCACACTGGAGTAATAGACCATCAAGGAACATCTCAGC 2580  
Db 3178 ATGGCGCTGGAGGACCGCGCCACACTGGAGTAATAGACCATCAAGGAACATCTCAGC 3237  
QY 2581 AGCAAGAGTCCCAACATATGGGTGACCTTGTGGAGAGCTTGGACAGCTGCGCCGCAAA 2640  
Db 3238 AGCAAGAGTCCCAACATATGGGTGACCTTGTGGAGAGCTTGGACAGCTGCGCCGCAAA 3297  
QY 2641 GTTCCACAGCGGAGCGCTCTCTGGGTCTCCCGGAGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 2700

Db 3296 GTTCCACAGCGGAGAGCGCTCCCTGGTCCCGGGAGCGCTCCCTGTCTCAGACCGGTCTA 3357  
QY 2701 AGCAAGCGCTGGAATGCACCACTCCTCTTCTTACGGGGTTSACTATAAGAGGAGCTAC 2760  
Db 3358 AGCAAGCGCTGGAATGCACCACTCCTCTTCTTACGGGGTTSACTATAAGAGGAGCTAC 3417  
QY 2761 CCACAGAACTCTGCTACAGAGAGCCACAGGCGCACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2820  
Db 3418 CCACAGAACTCTGCTACAGAGAGCCACAGGCGCACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3477  
QY 2821 TCCCTCAATTCCTCTCACTCTCCAGAAACACAGAGCTTTGGCAGGGAGAGCAACCGCGG 2880  
Db 3478 TCCCTCAATTCCTCTCACTCTCCAGAAACACAGAGCTTTGGCAGGGAGAGCAACCGCGG 3537  
QY 2881 CCGGCGCGGAGAGGCTGACTCCATCCAGGTGCACAGCTCCAGCCATCTCTGGCCAGGCC 2940  
Db 3538 CCGGCGCGGAGAGGCTGACTCCATCCAGGTGCACAGCTCCAGCCATCTCTGGCCAGGCC 3597  
QY 2941 GTCACTGTCTGAGGAGCGCCGCTCAACGCTTCACTCACTCACTCACTCACTCACTCACTCACTCA 3000  
Db 3598 GTCACTGTCTGAGGAGCGCCGCTCAACGCTTCACTCACTCACTCACTCACTCACTCACTCACTCA 3657  
QY 3001 AAGCTGACGCTCTGCTTAAAGCGGAGCTTAAAGCGGAGCTTAAAGCGGAGCTTAAAGCGGAGCTTAA 3060  
Db 3658 AAGCTGACGCTCTGCTTAAAGCGGAGCTTAAAGCGGAGCTTAAAGCGGAGCTTAAAGCGGAGCTTAA 3717  
QY 3061 ACATCAATGAAGCCCAATGAAGCTGATCAATAA 3093  
Db 3718 ACATCAATGAAGCCCAATGAAGCTGATCAATAA 3750

RESULT 2  
AAA93617  
ID AAA93617 standard; DNA; 1498 BP.  
XX  
AC AAA93617;  
XX  
XT 16-JAN-2001 (first entry)  
DE Human semaphorin protein-like splice variant SECX 2864933-1 DNA.  
XX SECX protein; human; secreted; membrane-associated; cancer;  
KW proliferation regulator; differentiation regulator; non-malignant tumour;  
KW immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;  
KW infection; inflammatory disorder; arthritis; hematopoietic disorder;  
KW skin disorder; cardiovascular disorder; atherosclerosis; restenosis;  
KW neurological disease; Alzheimer's disease; trauma; wounding;  
KW spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;  
KW anti-HIV; antiinflammatory; antiarthritic; antiarteriosclerotic;  
KW neuroprotective; vulnery; antiallergic; antimicrobial; cardiant;  
KW dermatological; gene therapy; ds.  
XX Homo sapiens.  
XX CS  
XX WO200053742-A2.  
XX  
XX 14-SEP-2000.  
XX  
XX Q9-MAR-2000; 2000WO-US06280.  
XX  
XX Q9-MAR-1999; 99US-0123667.  
PR Q8-MAR-2000; 2000US-0123667.  
XX  
XX (CURA-) CURAGEN CORP.  
XX Shinkets RA;  
XX  
XX WPI; 2000-594318/56.  
DR P-PSDB; AAB23030.  
XX  
XX Novel human membrane associated or secreted polypeptides and  
PT polynucleotides useful for diagnosis, prevention and treatment of  
PT pathological states such as cancer, immune, cardiovascular and

neurological disorders -

Claim 3; Fig 2; 151pp; English.

Sequences AAA93616-A194631 and AAA93673-A93676 represent nucleic acids which encode human SECX proteins (AA823029-823048). The SECX proteins of the invention are either secreted or membrane-associated proteins and act as regulator of cellular proliferation and differentiation. SECX proteins or nucleotides are useful for diagnosing the presence of, or predisposition to, a disease associated with altered levels of SECX proteins and nucleotides. The SECX proteins are also useful to screen compounds that modulate SECX activity or expression. The interaction of a SECX protein with other cellular proteins may be useful to modulate the activity of a partner protein, cellular proliferation, cellular differentiation and cell survival. SECX nucleotides are useful for the recombinant expression of SECX protein, and may be used to detect SECX mRNA or genetic lesions in the SECX gene. They may also be used to modulate SECX expression (e.g., using antisense oligonucleotides). SECX nucleic acid sequences are also useful for identifying a cell or tissue type in a biological sample, and in forensic biology. SECX primers or probes are useful for detecting the presence of SECX nucleotides and for screening tissue cultures for contamination. Diseases that may be treated or prevented using SECX proteins or nucleotides include cancer (e.g., colorectal carcinoma, prostate cancer), benign tumours, immune disorders (including autoimmune diseases, transplant rejection, allergies, AIDS), infections, inflammatory disorders, arthritis, haematopoietic disorders, skin disorders, cardiovascular disorders,therosclerosis, restenosis, neurological diseases (e.g., Alzheimer's disease), trauma (e.g., surgical or traumatic wounds, spinal cord injury), and skeletal disorders.

Sequence 3498 BP; 917 A; 966 C; 889 G; 725 T; 1 other;

```

Very Match          98.2%; Score 3038; DB 21; Length 3498;
1st Local Similarity 99.3%; Pred. NO. Q;
Matches 3072; Conservative 0; Mismatches 21; Indels 2; Gaps 2;

```

1 ATGAGGTCAGAAGCCTTGGTGGTATATTTCACACACGCGTACACTTTGGCTGGGCGCTGGTTTC 60

214 ATGAGGTCAGAGCCCTTGCTGCIATATTTCACACTGCTACACITTGCTGGGGCTGGTTTC 273

61 CCAGAAGATTCTGAGCCCAATCAGTATTTCGCATGGCAACTAACAACACAGTATCCGTC 120

274 CCAGAGATTCTGAGCCAAATCAGTATTGGCATGGCAAC...ATACAAACAGTATCCGGTG 333

121. TTTGTGGGCCACAAGCAGGACGGGAACACACACAGAGGCACAGGCTGGACATCCAGATG 180

334 TTTGTGGGGCCACAAGCCAGGACGGAAACACCCACACAGAGGCACAGGCTGGACATCCAGATG 393

181 ATATGATCATGAACGGGAACCCCTACATGCTGCTAGGGACCATATTTATACCTGTA 240

394 AATTATGATCATGAACCGAAGCCCTACATTCCTGCTAGGGACCATATTTATACTGTTGAT 453

241 ATAGACACATCACACACGGGAAGAAATTATTGTAGCAAAAACGACATGGAAATCTAGA 300

454 ATAGACACATCACACACGGAAAGAAATTATTTAGCAAAAACCTGACATGGAAACTAGA 513

301 CAGGGCCGATGTAGACACATGCGAGAAATGAAGGGGAAACATATAGGATGAGTGGCACAACCTT 360

514 CAGGCCGATGTAGACACATGCGAATGGAAGGCGAATACATATAGGATGATGACTGGCACAACTTT 573

301 ATATTAAGTCTCTCAATTAAGAAATACCAATGATGCATTGTTTGTCGTGGAACTAATGCC-TC 420

[illegible]

421 AACCCCTCCCTGCGAGTACATATAGATGGATACATGGATCCCATCGGGGATGAAACAGC 480

034 AACCCATTCCTGCGAATAACCATTAAGATGGATGGATACCATTTCAAGTGGAAATCAGG

[illegible]

V374 GGAAAGGCGATGGCCCATCTGTGTTGCCCCATTACATGCCCCAATCCTTGCACCTGTTGCAGATGGK 73

541 AAACATATACAGCCACAGTGACTCTCTTGCATGACCGAGTCAITTTACCGGAGT 600  
754 AAACATATACAGCCACAGTGACTCTCTTGCATGACCGAGTCAITTTACCGGAGT 813  
601 CTTGGAGAAAGCCCTACCTCCGACCGGTCAACGACGANTCAAATGSGTTGAAGAACA 660  
814 CTTGGAGAAAGCCCTACCTCCGACCGGTCAACGACGANTCAAATGSGTTGAAGAACA 873  
661 TACTTTTGTCAAGCCGTGGAATTACGGAGATTATATCTACTTCTCTTCAGGGAATAAGCA 720  
874 TACTTTTGTCAAGCCGTGGAATTACGGAGATTATATCTACTTCTCTTCAGGGAATAAGCA 933  
721 GTGAGATATACACCAATGGGAAGGTAGTTTCCCAAGAGTGGCTACAGGTTTGTGAAGAT 780  
934 GTGAGATATACACCAATGGGAAGGTAGTTTCCCAAGAGTGGCTACAGGTTTGTGAAGAT 993  
781 GATATGGAGAGATCTCAAGAGTCCCTGGAGAAACAGTGGACGCTCGTCTCGAAGCGCGC 840  
994 GATATGGAGAGATCTCAAGAGTCCCTGGAGAAACAGTGGACGCTCGTCTCGAAGCGCGC 1053  
841 TTGAACCTGCTCAGTTCTGTGAGACTCTCATTTTTATTTCACAAATCTCCAGGCAAGTTACA 900  
1054 TTGAACCTGCTCAGTTCTGTGAGACTCTCATTTTTATTTCACAAATCTCCAGGCAAGTTACA 1111  
901 GATGIGATTGATATCAAGGGGCGTGATGTCTCTGGCAAGCTTTCTACACCTTATAAC 960  
1114 GATGTGATTCGATATCAAGGGGCGTGATGTCTCTGGCAAGCTTTCTACACCTTATAAC 1173  
961 AGCATCCCTGGGTCCTGAGTCTGTGGCTATGACATGCTTGACATTCGCCAGTGTTTTACT 1020  
1174 AGCATCCCTGGGTCCTGAGTCTGTGGCTATGACATGCTTGACATTCGCCAGTGTTTTACT 1233  
1021 GGGAGATTCAAGGAACAGAAAGTCTCTGTATTCACGCTGGACACCAATCTCTGATGAACGA 1080  
1234 GGGAGATTCAAGGAACAGAAAGTCTCTGTATTCACGCTGGACACCAATCTCTGATGAACGA 1293  
1081 GTTCTTAAGCCAGGCGAGGTTGGTGTGCTGGGTCTCTCTTGAAGAGATATGCAACG 1140  
1294 GTTCTTAAGCCAGGCGAGGTTGGTGTGCTGGGTCTCTCTTGAAGAGATATGCAACG 1353  
1141 TCCAAATGAGTTCCTGTGATACCTCGAAGTTCATCAAGACGACACCGGCTCATGGATGAG 1200  
1354 TCCAAATGAGTTCCTGTGATACCTCGAAGTTCATCAAGACGACACCGGCTCATGGATGAG 1411  
1201 GCAGTGCCCTGCATCTTCAACAGGCCATGGTCTCTGAGACAAATGTCAGATACCGCCCT 1260  
1414 GCAGTGCCCTGCATCTTCAACAGGCCATGGTCTCTGAGACAAATGTCAGATACCGCCCT 1473  
1261 ACCAAAATTGCGAGTGGACAGAGTGTGGGCCATATCAGAAATCACACTGTGGTTTTCTG 1320  
1474 ACCAAAATTGCGAGTGGACAGAGTGTGGGCCATATCAGAAATCACACTGTGGTTTTCTG 1533  
1321 GGATCAGAGAAGGGAATCATCTTGAAGTTTGTGGCCAGAAATAGGAATATAGTGGTTTTCTTA 1380  
1534 GGATCAGAGAAGGGAATCATCTTGAAGTTTGTGGCCAGAAATAGGAATATAGTGGTTTTCTTA 1593  
1381 AATGACAGCCTTTTCTCGGAGGAGATCAGTGTTTTACAACCTCTGAAAAATTCACGCTATGAT 1440  
1594 AATGACAGCCTTTTCTCGGAGGAGATCAGTGTTTTACAACCTCTGAAAAATTCACGCTATGAT 1653  
1441 GGAGTCGAGACAAAAGSATCATGGGCATGACGCTGGACAGACCAAGCAGCTCTCTGTAT 1500  
1654 GGAGTCGAGACAAAAGSATCATGGGCATGACGCTGGACAGACCAAGCAGCTCTCTGTAT 1713  
1501 GTTGGGTTCTCACTGTGTGATAAAGGTTCCCGTTGGCCGGTGTGAACAGCATAGGGAAG 1560  
1714 GTTGGGTTCTCACTGTGTGATAAAGGTTCCCGTTGGCCGGTGTGAACAGCATAGGGAAG 1773  
1561 TGTAAARAAACCTGTATTGCCCTCAGAGACCCATTAATGTGGATGGATTAAGGAAGGTGGT 1620  
1774 TGTAAARAAACCTGTATTGCCCTCAGAGACCCATTAATGTGGATGGATTAAGGAAGGTGGT 1833  
1621 GCCTGCGACCCATTATCACCCAAACAGCAGACTGACTTTTGAGCAGGACCATAGAGCGTGGC 1680



1834	DB	GCCTCAGCCATTTATCACCACAGACAGACTGACCTTTTGACAGGACATAGAGCTGCG	1893
1881	QY	AATACAGATGGCTGGGGACATGTCACAAATTCCTTTGTGGCACTGAATGGGCAATTCAGAT	1740
1894	DB	AATACAGATGGCTGGGGACATGTCACAAATTCCTTTGTGGCACTGAATGGGCAATTCAGAT	1953
1741	QY	TCCTCTTTGCCAGACAAACACATCAGATTCGAGGGCTCAAGAGGGGTATAGTCTAGG	1800
1954	DB	TCCTCTTTGCCAGACAAACACATCAGATTCGAGGGCTCAAGAGGGGTATAGTCTAGG	2013
1801	QY	GGAGAAATGCTGGAGTGGAGCATTTGCTTGACTCAGCTGACAGACAGACGCTTTGGGG	1860
2014	DB	GGAGAAATGCTGGAGTGGAGCATTTGCTTGACTCAGCTGACAGACAGACGCTTTGGGG	2073
1861	QY	GGAGTGTCTCCCATTAATCACCAGACAAAGAGGGAGTGAATCGGGAAGATTACCTCAAA	1920
2074	DB	GGAGTGTCTCCCATTAATCACCAGACAAAGAGGGAGTGAATCGGGAAGATTACCTCAAA	2133
1921	QY	GGCCAGCACAGCTGGTTCGGTCAACCTCTTGGGCAATTCAGTCAATCCTGGCTTTGGCTC	1980
2134	DB	GGCCAGCACAGCTGGTTCGGTCAACCTCTTGGGCAATTCAGTCAATCCTGGCTTTGGCTC	2193
1981	QY	ATGGGGGCGGTCTTCGCGGCATCACCGTCTACTGGTCTGTGTAATCATCGGCGCAAGAC	2040
2194	DB	ATGGGGGCGGTCTTCGCGGCATCACCGTCTACTGGTCTGTGTAATCATCGGCGCAAGAC	2253
2041	QY	GTGGCTGTGGTCAGCGCAAGAGAGAGAGCTCACCCACTCGCGCGGGGCTCCATGAGC	2100
2254	DB	GTGGCTGTGGTCAGCGCAAGAGAGAGAGCTCACCCACTCGCGCGGGGCTCCATGAGC	2313
2101	QY	ACGGTCACCAAGCTCAGCGGCTCTTTTGGGGACACTCAATTCAAAGACCCAAAGCGGAG	2160
2314	DB	ACGGTCACCAAGCTCAGCGGCTCTTTTGGGGACACTCAATTCAAAGACCCAAAGCGGAG	2373
2161	QY	GGCATCCTCAGCGCATCATGCAACAGGCAAGCTCGCCACTTCGCGGCAACAGCGGCAAG	2220
2374	DB	GGCATCCTCAGCGCATCATGCAACAGGCAAGCTCGCCACTTCGCGGCAACAGCGGCAAG	2433
2221	QY	ATGCTCATTAAGACACACAGCAACCACTTGGAGCTCAGCGGCTTCGCCACCCACAGAGTCA	2280
2434	DB	ATGCTCATTAAGACACACAGCAACCACTTGGAGCTCAGCGGCTTCGCCACCCACAGAGTCA	2493
2281	QY	ACCCCAACGCTGCAGCAGAAACGGAAACCCAGCGCGGAGCGCGAGTGGGAGAGGAAAC	2340
2494	DB	ACCCCAACGCTGCAGCAGAAACGGAAACCCAGCGCGGAGCGCGAGTGGGAGAGGAAAC	2553
2341	QY	CAGAACCTCATCAATGCTGTGCAAAAGGACATCGCCGCCATGGGCTCGCTGTGATTGCG	2400
2554	DB	CAGAACCTCATCAATGCTGTGCAAAAGGACATCGCCGCCATGGGCTCGCTGTGATTGCG	2613
2401	QY	ACGGAGCTGCCCTGGGGGCTTCGCCAGCCATATGCCAGGCTGTGTCTGTGCCATTC	2460
2614	DB	ACGGAGCTGCCCTGGGGGCTTCGCCAGCCATATGCCAGGCTGTGTCTGTGCCATTC	2673
2461	QY	ACGCAGCGGCTACACAGATAGTAGTGGTGGACAGCGGCAAAATGACGAGGTGGCGGAG	2520
2674	DB	ACGCAGCGGCTACACAGATAGTAGTGGTGGACAGCGGCAAAATGACGAGGTGGCGGAG	2733
2521	QY	ATGGCGCTGGAGGACACAGGCGGCCACACTGGAGTATAGACCATCAAGGAACATCTCAGC	2580
2734	DB	ATGGCGCTGGAGGACACAGGCGGCCACACTGGAGTATAGACCATCAAGGAACATCTCAGC	2793
2581	QY	AGCAAGTCTCCAAACCATGGGTGAACCTTGTGGAACCTGACAGGCTGCGCCCAAA	2640
2794	DB	AGCAAGTCTCCAAACCATGGGTGAACCTTGTGGAACCTGACAGGCTGCGCCCAAA	2853
2641	QY	GTTCACAGCGGGAGGCTCCCTGGGTCCCGCGGGAGGCTCCCTGTCTCAGACCGGCTA	2700
2854	DB	GTTCACAGCGGGAGGCTCCCTGGGTCCCGCGGGAGGCTCCCTGTCTCAGACCGGCTA	2913
2701	QY	AGCAAGCGGTGAAATGCACACTCTCTTCTATCGGGGTTCACTATAAGAGAGCTAC	2760

2914	ACGACAGCGCTGGAAATGCACCACCTCCTCTTCCTACGGGGTTGACTATTAAGGAGGTAC	2878
QY	CCACGAACAATCGCTCAGGAGAAGCCACCAAGGCCACC-ACCTCTCAAAGAAAACAACTATAA	2819
DQ		
2974	CCACGAACAATCGCTCAGGAGAAGCCACCTGACCAACCTACTCTCATCAGAAGCAACACTAA	3033
QY	CHCCTTCCAATTCTCTTC-ACCTCTCAGAAACACAGAGCTTTGGCAGGGGAGACACACCGC	2878
DB		
3034	CCCCGACAAAT-CANCTCTTGACTTCAAAGGAGCACAGAGCTTTGGCAGGGGAGACACACCGC	3093
QY	CGCCCGCCGCCCGCAGAGAGGTGACTTCCATCCAGAGTGCACACACTCCCAGGCCATCTGGCCAGG	2938
DB		
3094	CGCCCGCCGCCCGCAGAGAGGTGACTTCCATCCAGAGTGCACACACTCCCAGGCCATCTGGCCAGG	3153
QY	CSGTGACTGTTCGAGGACAGCCAGCCAGCCCTAACGCCCTACAACCTCACTGACAAGGTCGGGGC	2998
DB		
3154	CSGTGACTGTTCGAGGACAGCCAGCCAGCCCTAACGCCCTACAACCTCACTGACAAGGTCGGGGC	3213
QY	TGAAGCGTACGCCCTTCGCTAAAGCGGACGCTAGACCCGCCCAAGACATCTTTGCTCCCTTT	3058
DB		
3214	TGAGGCGTAGGCCCTCGCTAAGCGCGGAGCTACCCCGAAGACCATCTTGTCTCCCTTT	3273
QY	CCACATTCATGAAGCCCANICATCGCTGTACATAA	3093
DB		
3274	CCACATTCATGAAGCCCANICATCGCTGTACATAA	3308
RESULT 3		
ABX71103		
ID	ABX71103 standard; cDNA: 4280 BP.	
XX	ABX71103;	
AC		
XX		
XX		
DT	05-MAR-2003 (first entry)	
XX		
DE	Novei human cDNA sequence #328.	
XX		
KW	Human; gene; ss; nervous system disorder; peripheral neuropathy;	
KW	Huntington's disease; amyotrophic lateral sclerosis; haemophilia;	
KW	neurodegenerative disease; Parkinson's disease; Alzheimer's disease;	
KW	autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;	
KW	insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;	
KW	ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;	
KW	fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;	
KW	coagulation disorder; cancer; tumour; inflammatory disease;	
KW	septic shock; Crohn's disease; anaphylaxis; proliferation; chemotactic;	
KW	differentiation; stem cell growth factor; haematopoiesis; chemokinetic;	
KW	haemostatic; antiinflammatory; expressed sequence tag; ESI.	
XX		
CS	Homo sapiens.	
XX		
PN	W0206281731-A2.	
XX		
PD	17-GCT-2002.	
XX		
PT	29-JAN-2002; 2002WO-US01222..	
XX		
PR	30-JAN-2001; 2001US-0774528.	
XX		
PA	{HYSE-} HYSEQ INC.	
FA	(GOOD/) GOODRICH R W.	
XX		
PI	Tang TY, Jiao C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;	
PI	Xue AZ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;	
XX		
DR	WPI: 2003-058563/05.	
XX		
PT	Novel polypeptide useful for treating neurodegenerative diseases,	
PT	miveld or lymphoid cell disorders, bone disorders, mechanical and	
PT	traumatic disorders, coagulation disorders, and inflammatory diseases	
PT	.	
PS	Claim 1: Page -; 612pp: English.	

Claim 1: Page -: 622pp; English.

	xx	This invention relates to the cDNA sequences encoding an isolated novel human polypeptide. The protein encoded by the nucleic acid o; the invention is useful for treating central and peripheral nervous system diseases (e.g. peripheral neuropathy, Huntington's disease, amyotrophic lateral sclerosis); neurodegenerative diseases (e.g. Parkinson's disease, Alzheimer's disease); autoimmune disease (e.g. systemic lupus erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus); myeloid or lymphoid cell disorders (e.g. anemia and thrombocytopoenia); wounds, burns; bone disorders (e.g. osteoporosis, osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head trauma); lung or liver fibrosis; reperfusion injury in various tissues; bacterial, viral or fungal infections; allergic conditions such as allergic rhinitis, asthma; coagulation disorders (e.g. hemophilia); cancer and tumors; and inflammatory diseases (e.g. septic shock, Crohn's disease, anaphylaxis). The protein may be used to inhibit the growth, infection or function of infectious agents such as bacteria, fungi, viruses, or to effect bodily characteristics, biorhythms or circadian cycles of rhythms. The protein may also have proliferation/differentiation, stem cell growth factor, haematopoiesis regulation, immune stimulating or suppressing, chemotactic/chemokinetic, haemostatic and thrombolytic, receptor/ligand, and antiinflammatory activities. The cDNA sequences of the invention are useful for expressing recombinant protein for analysis. The present sequence represents a novel human cDNA sequence of the invention. CC This sequence is an expressed sequence tag (EST) and was identified using subtractive hybridisation.
	xx	
	SQ	Sequence 4280 WP: 1165 A; 1165 C; 100 G; 925 T; 0 other;
		Query Match 98.0%; Score 3032; DB 25; Length 4280;
		Best Local Similarity 98.4%; Pred. No. 0;
		Matches 3093; Conservative 0 Mismatches 0 Indels 51; Gaps 1;
QY	1	ATGAGGTCAGAAACCTTGGTGCTAATTTCACACTGTACACTTTGTGGCGCTGGTTC 60
DB	267	ATGAGGTCAGAACCTTGGTGCTATATTCACATGSCACACTTGTGGCGCTGGTTC 326
QY	61	CAGAGAATCTTGAGCCCAATCAGTATTCGCAAGGCACACTATACAACACAGATCCCGTG 120
DB	327	CAGAGAATCTTGAGCCCAATCAGTATTCGCAAGGCACACTATACAACACAGATCCCGTG 386
QY	121	TTTGTGGGCCACAGCCAGGCGGACACCACACAGAGCCACAGCTGGACATCCAGNTG 180
DB	387	TTTGTGGGCCACAGCCAGGCGGACACCACACAGAGCCACAGCTGGACATCCAGAG 445
QY	181	ATTATGATCATGAAGGAAACCTTCACATTCCTGCTAGGACCATATTTATATCTTTGAT 240
DB	447	ATTATGATCATGAAGGAAACCTTCACATTCCTGCTAGGACCATATTTATCTGTTGAT 506
QY	241	AAGACACATCACACGAGAGAAATTTATGTACCAAAAACACTGCATGGAANTCTAGA 300
DB	507	ATAGACATATCACACGAGAGAAATTTATGTACCAAAAACACTGCATGGAANTCTAGA 566
QY	301	CAGGCGGATGTAGACACATGCAGATTAAGAGGAAAAACATAAGGATGAGTGCACAACTTT 360
DB	567	CAGGCGGATGTAGACACATGCAGATTAAGAGGAAAAACATAAGGATGAGTGCACAACTTT 626
QY	361	ATTAAAGTCTTCTAAGAGAAACAGATGATGATTTGTTGCTCTGGAACTAATGCGCTTC 420
DB	627	ATTAAAGTCTTCTAAGAGAAACAGATGATGATTTGTTGCTCTGGAACTAATGCGCTTC 686
QY	421	AACCCCTTCTTCAGAAACTATTAAGATGATACATTTGGAACCATTCGGSGATGAATTCAG 480
DB	687	AACCCCTTCTTCAGAAACTATTAAGATGATACATTTGGAACCATTCGGSGATGAATTCAG 746
QY	481	GGAATGGCCAGATGCCCATPATGATGCCAAACATGCCAACCTTGCACCTGTTTGCAGATGGA 540
DB	747	GGAATGGCCAGATGCCCATPATGATGCCAAACATGCCAACCTTGCACCTGTTTGCAGATGGA 806
QY	541	AAACTATATCAGCCACAGTAGTACTGACTTCCTTGTCCATGACGAGTGCATTTACCGGAGT 600
DB	807	AAACTATATCAGCCACAGTAGTACTGACTTCCTTGTCCATGACGAGTGCATTTACCGGAGT 866
QY	601	CTTGGAGAAAGCCCCTACCCCTGGGACCGCTCAAGCAGCATTCACAAATGTTTGAAGAACCA 666
DB	667	CTTGGAGAAAGCCCCTACCCCTGGGACCGCTCAAGCAGCATTCACAAATGTTTGAAGAACCA 926
QY	661	TACTTGTTCAGACCGCTGGATTACGAGAGATTATATCTACTTCTTCTTCAGAGGAAATAGCA 720
DB	927	TACTTGTTCAGACCGCTGGATTACGAGAGATTATATCTACTTCTTCTTCAGAGGAAATAGCA 986
QY	721	GTGGAGTATAACACCATTCGGGAAAGTAGTTTTCCCAAAGATGGCTCAGGTTTGTAAAGAT 780
DB	987	GTGGAGTATAACACCATTCGGGAAAGTAGTTTTCCCAAAGATGGCTCAGGTTTGTAAAGAT 1046
QY	781	GATATGGGAGGATCTCAAGAGATCTCTGGAGAAACAGTGGAGAGCTGTTCTCTAGAGCGCGC 840
DB	1047	GATATGGGAGGATCTCAAGAGATCTCTGGAGAAACAGTGGAGAGCTGTTCTCTAGAGCGCGC 1106
QY	841	TTGAACITGCTCAGTTCCTGGAGAGCTCTCATTTTATTTCACATTCCTCCAGGCGAGTTACA 900
DB	1107	TTGAACITGCTCAGTTCCTGGAGAGCTCTCATTTTATTTCACATTCCTCCAGGCGAGTTACA 1166
QY	901	GATGTGATTCGATCAACGGCGGTGATGTTCTCTGGCAAGCTTTTCTACACCTATATAAC 960
DB	1167	GATGTGATTCGATCAACGGCGGTGATGTTCTCTGGCAAGCTTTTCTACACCTATATAAC 1226
QY	961	AGCATCCCTGGGTCTGCACTCTGTGCCCTATGACATGCTTGACATTCGACATGCTTTTACT 1020
DB	1227	AGCATCCCTGGGTCTGCACTCTGTGCCCTATGACATGCTTGACATTCGACATGCTTTTACT 1286
QY	1021	GGGAGATTCAAGGAACAGAAATCTCTCTGATTCACCTGGACACAGTTCCTGATGAACGA 1080
DB	1287	GGGAGATTCAAGGAACAGAAATCTCTCTGATTCACCTGGACACAGTTCCTGATGAACGA 1346
QY	1081	GTTCCTTAAGCCAGCCAGCTGTGCTGCTGCTCATCTCTCTTASAAAGATATGCAACC 1140
DB	1347	GTTCCTTAAGCCAGCCAGCTGTGCTGCTGCTCATCTCTCTTASAAAGATATGCAACC 1406
QY	1141	TCCAATGAGTTCCCTGATGATACCTGAACTTCATCAAGAGCCACCGCTCATGGATGAG 1200
DB	1407	TCCAATGAGTTCCCTGATGATACCTGAACTTCATCAAGAGCCACCGCTCATGGATGAG 1466
QY	1201	GCAGTGGCCCTCATCTTCAACAGGCCATGTTCTCTGAGAAACATNGSTTCAGATACCGCTT 1260
DB	1467	GCAGTGGCCCTCATCTTCAACAGGCCATGTTCTCTGAGAAACATNGSTTCAGATACCGCTT 1326
QY	1261	ACCAAAATTCGCACTGGACACAGCTCTGGGCAATCAASAATCACACTGTGTTTCTG 1320
DB	1527	ACCAAAATTCGCACTGGACACAGCTCTGGGCAATCAASAATCACACTGTGTTTCTG 1586
QY	1321	GCATCAGAGAGGAAATCATCTTGAAGTTTTTGCCAGATAGGAATAGTGGTTTCTTA 1380
DB	1587	GCATCAGAGAGGAAATCATCTTGAAGTTTTTGCCAGATAGGAATAGTGGTTTCTTA 1646
QY	1381	AATGACAGCTTTTCTCTGGAGAGATGATGTTTACAACTCTGAAAATGCAGCATATGAT 1440
DB	1647	AATGACAGCTTTTCTCTGGAGAGATGATGTTTACAACTCTGAAAATGCAGCATATGAT 1706
QY	1441	GGAGTCGAGAGCAAAAGGATCATGGGCATGAGTGGACAGAGCAAGCAGCTCTCTCTAT 1500
DB	1707	GGAGTCGAGAGCAAAAGGATCATGGGCATGAGTGGACAGAGCAAGCAGCTCTCTCTAT 1766
QY	1501	GTTCGGTCTCTTACCTGTGTATAAGGTTCCCTTGGCCGCTGGAACACATGGGAAG 1560
DB	1767	GTTCGGTCTCTTACCTGTGTATAAGGTTCCCTTGGCCGCTGGAACACATGGGAAG 1826
QY	1561	TGTAAAAAACCTGTATTTCCTCCACAGACCCATATTGTGATGGATTAAGGAAGGTGGT 1620
DB	1827	TGTAAAAAACCTGTATTTCCTCCACAGACCCATATTGTGATGGATTAAGGAAGGTGGT 1886
QY	1621	GCTCAGCCATTTATACCCCAACAGCAGATGACTTTTIGAGCAGGACATAGACGCTGGC 1680
DB	1887	GCTCAGCCATTTATATCCCCAACAGCAGATGACTTTTIGAGCAGGACATAGACGCTGGC 1946

QY	1681	AATACAGATGGICTGGGGGACTGTCAAAATTCCTTTTGGCACTGA-----	1721
Db	1947	AATACAGATGGICTGGGGGACTGTCAAAATTCCTTTTGGCACTGAATGACATTTCAACI	2006
QY	1727	-----ATGGGCAATTCAGTTCCCTCTTG	1749
Db	2007	CTCTTACCAGATAATGAATGCTTACAACACAGTGTATGGGCATTCAGTTCCCTCTTG	2066
QY	1750	CCGAGCACACCAACACATCAGATTCGAGGGCTCAAGAGGGGTATCAGTCTAGGGGAGGAATG	1809
Db	2067	CCGAGCACACCAACATCAGATTCGAGGGCTCAAGAGGGGTATCAGTCTAGGGGAGGAATG	2126
QY	1810	CTGACTGGAGACATCTGCTTGACTCACTGTGACAGACACACACCTTTTGGGGGAGTGTCT	1869
Db	2127	CTGACTGGAGACATCTGCTTGACTCACTGTGACAGACACACACCTTTTGGGGGAGTGTCT	2186
QY	1870	TCCCATTAATCACCACACAGAAAGAGGAGTGATTCGGGAAAGTTACGCTCAAAAGCCACAGAC	1929
Db	2187	TCCCAANTCACCACAGAAAGAGGAGTGATTCGGGAAAGTTACGCTCAAAAGCCACAGAC	2244
QY	1930	CAGCTGTTCCCGTCACCCCTCTTGGGCATTCGAGTGCATCTGCTTTCGTCATGSGGGGC	1989
Db	2247	CAGCTGTTCCCGTCACCCCTCTTGGGCATTCGAGTGCATCTGCTTTCGTCATGSGGGGC	2306
QY	1990	GTCTCTCGGGCATCAGGCTTACTGGGTCTGTGCATCATCGCGCAAGACGTGGTGTG	2049
Db	2307	GTCTCTCGGGCATCAGGCTTACTGGGTCTGTGCATCATCGCGCGCAAGACGTGGTGTG	2366
QY	2050	GTCCAGCGCAAGAGAGAGGAGTCACCCACTCGCGGGGTCCCATGAGCACGCTCAC	2109
Db	2367	GTCCAGCGCAAGAGAGAGGAGTCACCCACTCGCGGGGTCCCATGAGCACGCTCAC	2426
QY	2110	AAGCTCAGCGGCTCTTTGGGACACTCANTCCAAAGACCCAAAGCGGAGGCGCATCTC	2169
Db	2427	AAGCTCAGCGGCTCTTTGGGACACTCANTCCAAAGACCCAAAGCGGAGGCGCATCTC	2486
QY	2170	ACGCCACTCATGACAAACGCAAGCTGGACACTCCGCGCAACACGCGGAGGCGCATCTC	2229
Db	2487	ACGCCACTCATGACAAACGCAAGCTGGACACTCCGCGCAACACGCGGAGGCGCATCTC	2546
QY	2230	AAGCAGACAGCACCACTGGACCTTGACGGGCTTCGCCACCGCACGACGATCAACG	2289
Db	2547	AAGCAGACAGCACCACTGGACCTTGACGGGCTTCGCCACCGCACGAGTCAACG	2606
QY	2290	CTGCAGCAGAAAGCGGAGGCCAGCGCGGCGGAGCGGAGTGGAGAGGAACCAAGACCTC	2349
Db	2607	CTGCAGCAGAAAGCGGAGGCCAGCGCGGCGGAGTGGAGAGGAACCAAGACCTC	2666
QY	2350	ATCAATGCCCTGCACAAAGACATTCGCCCCCATGGCTTCCCTCTGTATTCACGGACCTG	2409
Db	2667	ATCAATGCCCTGCACAAAGACATTCGCCCCCATGGCTTCCCTCTGTATTCACGGACCTG	2726
QY	2410	CCCTCGGGGCTTCCGCCACACATTCGCCAGCGTGGTGTCTCTGCCCATCAGCGACGAC	2469
Db	2727	CCCTCGGGGCTTCCGCCACACATTCGCCAGCGTGGTGTCTCTGCCCATCAGCGACGAC	2786
QY	2470	GGCTACAGCATCAGTACGTTGACAGACCCCAAATGAGCGAGGTGGCCACAGATGGCGCTG	2529
Db	2787	GGCTACAGCATCAGTACGTTGACAGACCCCAAATGAGCGAGGTGGCCACAGATGGCGCTG	2846
QY	2530	GAGGACAGCGCCGACACTCGAGTATAAGACCATCAAGGAACATCTCAGCAGCAAGAGT	2589
Db	2847	GAGGACAGCGCCGACACTCGAGTATAAGACCATCAAGGAACATCTCAGCAGCAAGAGT	2906
QY	2590	CCCAACCATGGGTGAACCTTGTGGAGACCTGGACAGCCTGCCACAGATGGCGCTG	2649
Db	2907	CCCAACCATGGGTGAACCTTGTGGAGACCTGGACAGCCTGCCACAGATGGCGCTG	2966
QY	2650	CGGAGACCTTCCCTGGGTCCCGCGGAGCCTCCCTCTCAGACCGGTCTAAGCAAGCGG	2709
Db	2967	CGGAGACCTTCCCTGGGTCCCGCGGAGCCTCCCTCTCAGACCGGTCTAAGCAAGCGG	3026
QY	2710	CTGAAATGCACCACCTCCTCTTCTCAGCGGGTGTGACTATAGAGAGACCTACCCACGAAC	2769

Db	3027	CTGAAATGACCACTCTCTCTTCTACGGGGTTGACTATAAGAGGAGCTACCCACGAAC	3086
Qy	2770	TCGCTCAGGAGGACCCAGGCGACCACTCTCAAAAGAACACACTAACTCCTCCAAT	2829
Db	3087	TCGCTCAGGAGGACCCAGGCGACCACTCTCAAAAGAACACACTAACTCCTCCAAT	3146
Qy	2830	TCCTCTCACCTCTCTCAGAAACAGAGCTTTGGCAGGGGAGACACACCGCGCGCGCCCG	2889
Db	3147	TCCTCTCACCTCTCTCAGAAACAGAGCTTTGGCAGGGGAGACACACCGCGCGCGCCCG	3206
Qy	2890	CAGAGGAGTGAATCCATCCAGGTGCAGACTCCAGCCATCTGGCCAGGCGGTGACTGTC	2949
Db	3207	CAGAGGAGTGAATCCATCCAGGTGCAGACTCCAGCCATCTGGCCAGGCGGTGACTGTC	3266
Qy	2950	TCGAGGAGGAGCCAGCCCTCAAGCCCTACAACTCACTGACAAAGGTCGGGGCTGAAGCGTACG	3009
Db	3267	TCGAGGAGGAGCCAGCCCTCAAGCCCTACAACTCACTGACAAAGGTCGGGGCTGAAGCGTACG	3326
Qy	3010	CCCTCGGTAAAGCGGGAGGTACCCCGCAACCATCTTTCCTCCCTTTCCACATCCATG	3069
Db	3327	CCCTCGGTAAAGCGGGAGGTACCCCGCAACCATCTTTCCTCCCTTTCCACATCCATG	3386
Qy	3070	AAGCCCAATGATGGGTGTACATAA	3093
Db	3387	AAGCCCAATGATGGGTGTACATAA	3410
RESULT 4			
AAH18729			
IT	AAH18729 standard; cDNA: 6060 BP.		
XX	AAH18729;		
AC	AAH18729;		
DT	26-JUN-2001 (first entry)		
XX	Human cDNA sequence SEQ ID NO:19004.		
DE	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.		
KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.		
XX	OS: Homo sapiens.		
XX	EP1074617-A2.		
XX	07-FEB-2001.		
XX	28-JUL-2000; 2000EP-0116126.		
XX	29-JUL-1999; 99JP-0248036.		
PR	27-AUG-1999; 99JP-0300253.		
PR	11-JAN-2000; 2000JP-0118776.		
PR	02-MAY-2000; 2000JP-0183767.		
PR	09-JUN-2000; 2000JP-0241899.		
XX	(HELI-) HELIX RES INST.		
XX	Ota T, Isogai I, Nishikawa T, Hayashi K, Saito K, Yamamoto J;		
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki I;		
XX	WPI: 2001-318749/34.		
DR	Primer sets for synthesizing polynucleotides, particularly the 5602		
XX	full-length cDNAs defined in the specification, and for the detection		
PT	and/or diagnosis of the abnormality of the proteins encoded by the		
PT	full-length cDNAs.		
XX	Claim 8; SEQ ID 19004; 2537pp + CD ROM; English.		
PS	The present invention describes primer sets for synthesizing 5602		
XX	full-length cDNAs defined in the specification. Where a primer set		
CC	comprises: (a) an oligo-dr primer and an oligonucleotide complementary		
CC	to the complementary strand of a polynucleotide which comprises one of		
CC	the 5602 nucleotide sequences defined in the specification where the		

CC oligonucleotide comprises at least 15 nucleotides; or (2) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03156 to AAH3628 and  
CC AAH3633 to AAH18742 represent human cDNA sequences; AAH92446 to  
CC AAH3583 represent human amino acid sequences; and AAH3629 to AAH3632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
SQ Sequence 5060 BP; 1639 A; 1525 C; 1363 G; 1532 T; 0 other;

Query Match 96.0%; Score 2969.6; DB 22; Length 6060;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2994; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 96 CAACATATACAAACAGATCGCGTGTGTGGGCCAACAGCCAGCAGCAACCCACACA 155  
DB 1 CAACATATACAAACAGATATCGCGTGTGTGGGCCAACAGCCAGCAGCAACCCACACA 60

QY 156 GAGGCACAGGCTGACATCCAGATGATATGATCATGATCAAGGACACCTCTACATGCTGC 215  
DB 61 GAGGCACAGGCTGACATCCAGATGATATGATCATGATCAAGGACACCTCTACATGCTGC 220

QY 216 TAGGACCATATTTATCTGTGTATATAGACATACACAGCAACCAAAATTTATCTAG 275  
DB 12: TAGGACCATATTTATCTGTGTATATAGACATACACAGCAACCAAAATTTATCTAG 180

QY 276 CAAAAAAGTACATGGAATCTAGACAGGCGCATGTAGACACATGCAAGATGAAGGAAA 335  
DB 181 CAAAAAAGTACATGGAATCTAGACAGGCGCATGTAGACACATGCAAGATGAAGGAAA 240

QY 336 ACATAGGATGATGCCACACTTTATTAAGTCTCTTAAGAAACCAACGATGATGATT 395  
DB 241 ACATAGGATGATGCCACACTTTATTAAGTCTCTTAAGAAACCAACGATGATGATT 300

QY 396 GTTGTCTGTGGAAGTAAATGCTTCAAGCTTCTGTCAGAACTATAAGATGGAATGATT 455  
DB 301 GTTGTCTGTGGAAGTAAATGCTTCAAGCTTCTGTCAGAACTATAAGATGGAATGATT 360

QY 456 GGAACCATTCGGGATGAATTCAGCGGATGCGGATGCGCATATGATGCCAACAATCC 515  
DB 361 GGAACCATTCGGGATGAATTCAGCGGATGCGGATGCGCATATGATGCCAACAATCC 420

QY 516 CAACGTTGCAGCTGTTGACAGTGGAAACTATCTCAGCTACAGTGAATGCTTCCTGTC 575  
DB 421 CAACGTTGCAGCTGTTGACAGTGGAAACTATCTCAGCTACAGTGAATGCTTCCTGTC 480

QY 576 CATTCAGCAGTCAATTTACGAGTCTTGGAGAAAGCTTACCTGCGGACCGTCAAGCA 635  
DB 481 CATTCAGCAGTCAATTTACGAGTCTTGGAGAAAGCTTACCTGCGGACCGTCAAGCA 540

QY 636 CCATTCAAAATGGTTGAAGAACCATATCTTCTCAAGCCGCGGATTACGAGATTATAT 695  
DB 541 CCATTCAAAATGGTTGAAGAACCATATCTTCTCAAGCCGCGGATTACGAGATTATAT 600

QY 696 CTACTCTTCTTCAGGGAATAGAGTGGAGTATAACACCATGGGAAAGTACTGCTTCCC 755  
DB 601 CTACTCTTCTTCAGGGAATAGAGTGGAGTATAACACCATGGGAAAGTACTGCTTCCC 660

QY 756 AAGATGGCTCAGGTTGTGAAGATGATGGAGGATCTCAAGAGTCTCTGGAGAAACA 815  
DB 661 AAGATGGCTCAGGTTGTGAAGATGATGGAGGATCTCAAGAGTCTCTGGAGAAACA 720

QY 816 GTGAGGCTGCTCCTGAAGGCGCGCTTCAACTGCTCAGTCTCTCGAGACTCTCATTTA 875

DB 721 GTGAGCTGCTGCTGAGGCGGCTTAACTGCTCAGTTCCTGGAGACTCTCATTTTA 780  
QY 876 TTTCAACATTTCTCAGGCAATACAGATGTGATTCCTATCAAGGGGCGTGAITGTCCT 935  
DB 751 TTTCAACATTTCTCAGGCAATACAGATGTGATTCCTATCAAGGGGCGTGAITGTCCT 840  
QY 936 GGCACAGTTTCTACACTTATACACAGATCCCTGGGTCTGCGAGTCTGCGCTATGACAT 995  
DB 841 GGCACAGTTTCTACACTTATACACAGATCCCTGGGTCTGCGAGTCTGCGCTATGACAT 900  
QY 996 GCTTGACATTCGACAGTGTCTTACTGGAGATTCACGGAACAGAAAGTCTCTGATTCCAC 1055  
DB 901 GCTTGACATTCGACAGTGTCTTACTGGAGATTCACGGAACAGAAAGTCTCTGATTCCAC 960  
QY 1056 CTGGACACAGTTCTGATGAACAGATTCCTAAGCCAGGCCAGGTTGCTGTGCTGCTC 1115  
DB 961 CTGGACACAGTTCTGATGAACAGATTCCTAAGCCAGGCCAGGTTGCTGTGCTGCTC 1020

QY 1115 AACTCTCTTACAAAGATGACAGCTCCAATGAGTTCCTGCTGATGATACCTGGAATTCAT 1175  
DB 1021 AACTCTCTTACAAAGATGATGACAGCTCCAATGAGTTCCTGCTGATGATACCTGGAATTCAT 1080

QY 1176 CAAGACGACCGGCTCATGATGAGGAGTSCCTCTCCATCTTCAACAGGCCATGGTTCCT 1235  
DB 1081 CAAGACGACCGGCTCATGATGAGGAGTSCCTCTCCATCTTCAACAGGCCATGGTTCCT 1140

QY 1236 GAGAACATGATCAGATACCGGCTTACCAAAATTCAGATGACAGACAGTCTCTGGCATA 1295  
DB 1141 GAGAACATGATCAGATACCGGCTTACCAAAATTCAGATGACAGACAGTCTCTGGCATA 1200

QY 1296 TCAGAAATCAGACTGCTGTTTCTGGGATCAGAGAGGAAATCATCTTGAAGTTTITGGC 1355  
DB 1201 TCAGAAATCAGACTGCTGTTTCTGGGATCAGAGAGGAAATCATCTTGAAGTTTITGGC 1260

QY 1356 CAGAAATAGAAATAGTGTGTTTCTAAATGACAGGCTTCTCTGAGGAGATGAGTGTGTTA 1415  
DB 1261 CAGAAATAGAAATAGTGTGTTTCTAAATGACAGGCTTCTCTGAGGAGATGAGTGTGTTA 1320

QY 1416 CAACCTCTG-AAAAATGCACTATGATGAGTCTGAAGCAAAAGATCATGGGCATGACG 1474  
DB 1321 CAACCTCTGAAAAATGCACTATGATGAGTCTGAAGCAAAAGATCATGGGCATGACG 1380

QY 1475 TGACAGACAGCAGCAGCTCTCTGATGTTGGTCTCTCTACCTGTGTATAAAGTTCCCC 1534  
DB 1381 TGACAGACAGCAGCAGCTCTCTGATGTTGGTCTCTCTACCTGTGTATAAAGTTCCCC 1440

QY 1535 CTGGCGGCTGTGAACCACTGAGAGTCT-AAAAAACCCTGATTGCTCTCAGAGACCCA 1593  
DB 1441 CTGGCGGCTGTGAACCACTGAGAGTCTAAAAAACCCTGATTGCTCTCAGAGACCCA 1500

QY 1594 TATTGTGATGATAAAGGAGTGTGCTGCTGAGCCATTTATCACCACACAGCAGACTG 1653  
DB 1501 TATTGTGATGATAAAGGAGTGTGCTGCTGAGCCATTTATCACCACACAGCAGACTG 1560

QY 1654 ACTTTGACGAGCAGATAGAGGTTGGCAATACAGATGCTCTGGGGAGCTGTCAAAATCC 1713  
DB 1561 ACTTTGACGAGCAGATAGAGGTTGGCAATACAGATGCTCTGGGGAGCTGTCAAAATCC 1620

QY 1714 TTTGTGGCACTGAATGGGCAATTCAGTTCCCTCTTGGCCAGCAACACACATCAGATTG 1773  
DB 1621 TTTGTGGCACTGAATGGGCAATTCAGTTCCCTCTTGGCCAGCAACACACATCAGATTG 1680

QY 1774 ACGGCTCAAGAGGGTATGAGTCTAGGGGAGGATGCTGGAGTGGAGCATCTCTCTGAC 1833  
DB 1681 ACGGCTCAAGAGGGTATGAGTCTAGGGGAGGATGCTGGAGTGGAGCATCTCTCTGAC 1740

QY 1834 TCACCTGACAGCACAGACCTTTTGGGGCAGTGTCTTCCCATATATCACAAGACAAG 1893  
DB 1741 TCACCTGACAGCACAGACCTTTTGGGGCAGTGTCTTCCCATATATCACAAGACAAG 1800

QY 1894 GGAGTGAATTCGGGAAGTTACTCAAGGCCAGCAGCTGTTGCTCGCTCACCCTCTTG 1953  
DB 1741 GGAGTGAATTCGGGAAGTTACTCAAGGCCAGCAGCTGTTGCTCGCTCACCCTCTTG 1953



QY 61 CCAGAGANTCTGAGCCAAACGATTTTCGCAATGGCAACTATACAAACAGTAGTCCGGTG 120  
DB 186 CCAGAGANTCTGAGCCAAACGATTTTCGCAATGGCAACTATACAAACAGTAGTCCGGTG 245  
QY 121 TTTGTGGCCACAAGCCAGAGCAGAACACACACACAGAGCAGAGCGGTGACATCCAGATG 180  
DB 246 TTTGTGGCCACAAGCAGAGCAGAACACACACACAGAGCAGAGCGGTGACATCCAGATG 325  
QY 181 ATTAATGATCATGAAGGAGAACGTCATGATTTGCTGCTAGGAGACATATTTATATGTTGAT 240  
DB 306 ATTAATGATCATGAAGGAGAACGTCATGATTTGCTGCTAGGAGACATATTTATATGTTGAT 365  
QY 241 ATAGACATCATACACACAGGAGAGAAATTTATTTGAGCAAAAACATGACATGAAATGTAGA 300  
DB 366 ATAGACATCATACACACAGGAGAGAAATTTATTTGAGCAAAAACATGACATGAAATGTAGA 425  
QY 301 CAGGCCATGTAGACACATGAGAAATGAAGGAGAAACATAGAGATGAGTCCCAACATTTT 360  
DB 426 CAGGCCATGTAGACACATGAGAAATGAAGGAGAAACATAGAGATGAGTCCCAACATTTT 485  
QY 361 ATTAAGATTTCTTAAGAAATGAGGATGATGATTTGCTGCTGAGAACTATAGCTTC 420  
DB 486 ATTAAGATTTCTTAAGAAATGAGGATGATGATTTGCTGCTGAGAACTATAGCTTC 545  
QY 421 AACCTTCTCCAGAAACTATAAGATGATACATATGGAACCATTCGGGGATGAAATTCAGC 480  
DB 546 AACCTTCTCCAGAAACTATAAGATGATACATATGGAACCATTCGGGGATGAAATTCAGC 605  
QY 481 GGAATGGCCAGATGCCCATATGATGCCAAACATGCCAAGCTTGGACPTTTTGAGATGGA 540  
DB 606 GGAATGGCCAGATGCCCATATGATGCCAAACATGCCAAGCTTGGACPTTTTGAGATGGA 665  
QY 541 AAATTAATCTCAGCCACAGTACTGACTTCTCTGCGAATTCAGCCAGTCAATTCACCGAGT 600  
DB 666 AAATTAATCTCAGCCACAGTACTGACTTCTCTGCGAATTCAGCCAGTCAATTCACCGAGT 725  
QY 601 GTTGAGAAACCCATACCTCGACCGACCGTCAAGCAGCATTCANAAATGTTGAAAGAACCA 660  
DB 726 GTTGAGAAACCCATACCTCGACCGACCGTCAAGCAGCATTCANAAATGTTGAAAGAACCA 785  
QY 661 TACTTTGTTCAGCCGTGGATACGGAGATATATCTACITCTCTTCAGGGAATAGCA 720  
DB 786 TACTTTGTTCAGCCGTGGATACGGAGATATATCTACITCTCTTCAGGGAATAGCA 845  
QY 721 GTTGAGATATAACCATGGGAAGGTAGTTTTCGCAAGAGTGGCTCAGGTTTGTAGAAAT 780  
DB 846 GTTGAGATATAACCATGGGAAGGTAGTTTTCGCAAGAGTGGCTCAGGTTTGTAGAAAT 905  
QY 781 GNTATGGAGAGATCTCAAGAGTCTCGAGAAACAGTGGAGTCTGCTTCAGAGCGCGC 840  
DB 906 GAATGGAGAGATCTCAAGAGTCTCGAGAAACAGTGGAGTCTGCTTCAGAGCGCGC 965  
QY 841 TTGAAGTCTCAGTTCTCGAGAGTCTCATTTTATTTCAACATTTCTCCAGGAGTTTACA 900  
DB 966 TTGAAGTCTCAGTTCTCGAGAGTCTCATTTTATTTCAACATTTCTCCAGGAGTTTACA 1025  
QY 901 GAATGATTCGATACAGCGCGGTGATGTTGCTCGTGGCAAGGTTTCTACACCTTATAAC 960  
DB 1026 GAATGATTCGATACAGCGCGGTGATGTTGCTCGTGGCAAGGTTTCTACACCTTATAAC 1085  
QY 961 ASCATCCCTGGGTGCGAGTCTGCGCTACACATGCTTGACATTCGCACTGTTTACT 1020  
DB 1086 ASCATCCCTGGGTGCGAGTCTGCGCTATGACATGCTTGACATTCGCACTGTTTACT 1145  
QY 1021 GGGAGATTCAGGAACAGAACTCTCTGATTCACCTGGACACACAGTTCCTGATGAACGA 1080  
DB 1146 GGGAGATTCAGGAACAGAACTCTCTGATTCACCTGGACACACAGTTCCTGATGAACGA 1205  
QY 1081 GTTCTTAAGCCAGCGCAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
DB 1206 GTTCTTAAGCCAGCGCAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1265

QY 1141 TCCATAGAGTTCCTCTGATGATACCTGAACCTCATCAAGACGACCCGCTCATGATGAG 1200  
DB 1266 TCCATAGAGTTCCTCTGATGATACCTGAACCTCATCAAGACGACCCGCTCATGATGAG 1325  
QY 1201 GCAGTGGCCTCCATCTTCAACAGAGCCATGGTTCCTGAGAACTATGTCAGATACCGCTT 1260  
DB 1326 GCAGTGGCCTCCATCTTCAACAGAGCCATGGTTCCTGAGAACTATGTCAGATACCGCTT 1385  
QY 1261 ACCAAATATGAGTGGACACAGCTGCTGGGCCATATCAGAAATCAGATGCTGTTTCTG 1320  
DB 1386 ACCAAATATGAGTGGACACAGCTGCTGGGCCATATCAGAAATCAGATGCTGTTTCTG 1445  
QY 1321 GGATCAGAGAGAGGATCATCTTGAAGTTTGTGCGCAGAAATAGTGGTTTCTTA 1380  
DB 1446 GGATCAGAGAGAGGATCATCTTGAAGTTTGTGCGCAGAAATAGTGGTTTCTTA 1505  
QY 1381 AATGACAGCCTTTTCTGGAGAGATGATGTTTACAACITCTGAAAATGCAGTATGAT 1440  
DB 1506 AATGACAGCCTTTTCTGGAGAGATGATGTTTACAACITCTGAAAATGCAGTATGAT 1565  
QY 1441 GGATGCGAGACAAAGATCATGGGCATGACAGCTGGACAGCAGCAAGCAGTCTCTGTAT 1500  
DB 1566 GGATGCGAGACAAAGATCATGGGCATGACAGCTGGACAGCAGCAAGCAGTCTCTGTAT 1625  
QY 1501 GTTGGCTTCTCTACCTGTGTATAAAGCTTCCCTTGGCCGCTGTGAACGACATGGGAAG 1560  
DB 1526 GTTGGCTTCTCTACCTGTGTATAAAGCTTCCCTTGGCCGCTGTGAACGACATGGGAAG 1685  
QY 1561 TGTAAAAAACCCTGTTATGCTCCAGACCCCATATTTGTGGATGATTAAGGAAGTGGT 1620  
DB 1686 TGTAAAAAACCCTGTTATGCTCCAGACCCCATATTTGTGGATGATTAAGGAAGTGGT 1745  
QY 1621 GCCTCGACCATTTATCACCAACACAGACACTGACTTTTGAAGCAGGACATAGAGCGTGGC 1680  
DB 1746 GCCTCGACCATTTATCACCAACACAGACACTGACTTTTGAAGCAGGACATAGAGCGTGGC 1805  
QY 1681 AATACAGATGGTCTGGGAGCTGTCCAAATTCCTTTGTGGCAGCTGAATGGGCATTTCCAGT 1740  
DB 1806 AATACAGATGGTCTGGGAGCTGTCCAAATTCCTTTGTGGCAGCTGAATGGGCATTTCCAGT 1853  
QY 1741 TCCCTCTTGGCCAGCACCAACACATCAGATTCGACGGCTCAAGAGGGGTATGATCTAGG 1800  
DB 1854 TCCCTCTTGGCCAGCACCAACACATCAGATTCGACGGCTCAAGAGGGGTATGATCTAGG 1853  
QY 1801 GGAGGAATGCTGGACTGGAAGCATCTGCTTGACTACCTGACAGCAGACAGCCCTTTGGG 1860  
DB 1854 GGAGGAATGCTGGACTGGAAGCATCTGCTTGACTACCTGACAGCAGACAGCCCTTTGGG 1853  
QY 1861 GCAGTGTCTTCCCATATACCAACAGACAGAGGAGTGAATCGGGAAGTTACTCAAA 1920  
DB 1854 GCAGTGTCTTCCCATATACCAACAGACAGAGGAGTGAATCGGGAAGTTACTCAAA 1880  
QY 1921 GGCCAGCAGCAGCTGTTTCCCTCACCCTCTTGGCCATTCAGCTATCTGGCTTTCGTC 1980  
DB 1881 GGCCAGCAGCAGCTGTTTCCCTCACCCTCTTGGCCATTCAGCTATCTGGCTTTCGTC 1940  
QY 1981 ATGGGGGCGCTTCTCGGGCATCACCGTCTACTGCTGTGTGATCATCGCGCAAGAC 2040  
DB 1941 ATGGGGGCGCTTCTCGGGCATCACCGTCTACTGCTGTGTGATCATCGCGCAAGAC 2000  
QY 2041 GTGGGTGTGTCGCGCAAGGAGAGGAGTCAACCCACTCGCCGCGGGCTCCATGAGC 2100  
DB 2001 GTGGGTGTGTCGCGCGCAGAGAGAGGAGTCAACCCACTCGCCGCGGGCTCCATGAGC 2060  
QY 2101 AGCGTCACCAAGCTCAGCGGCTCTTTGGGACACTCAATCCAAAGACCCAAAGCCGAG 2160  
DB 2061 AGCGTCACCAAGCTCAGCGGCTCTTTGGGACACTCAATCCAAAGACCCAAAGCCGAG 2120  
QY 2161 GGCATCTCTCACGCCACTATSCACAGCGCAGCTCGCCACTCCCGGCAACACCGCCAG 2220  
DB 2121 GGCATCTCTCACGCCACTATSCACAGCGCAGCTCGCCACTCCCGGCAACACCGCCAG 2180  
QY 2221 ATGCTCATTAAGAGAGACCCAGCACCCACTGGACCTGACGGCTCCCTCCACCCAGAGTCA 2280

```

Db 2181 ATGTCATTAAGCAGACACACACACTGACGGCCCTCCGCCACCCAGAGTCA 2240
QY 2281 ACCCAAACGCTGCAGAGAGCGGAAGCCCGAGCGCGAGTGGGAGAGAAC 2340
Db 2241 ACCCAAACGCTGCAGAGAGCGGAAGCCCGAGCGCGAGTGGGAGAGAAC 2300
QY 2341 CAGAACCTCATCAATGCTGCACAAAGACATGCCGCCATGGGCTCCCTGTATTCC 2400
Db 2301 CAGAACCTCATCAATGCTGCACAAAGACATGCCGCCATGGGCTCCCTGTATTCC 2360
QY 2401 ACAGACCTGCCCTTGGGGCTGCCCGAGCACAATGCCAGCGTGGTCTCTGCCATC 2460
Db 2361 ACAGACCTGCCCTTGGGGCTGCCCGAGCACAATGCCAGCGTGGTCTCTGCCATC 2420
QY 2461 ACAGAGAGGGGTACAGACAATGATGAGTGGACAGCCCAAAATAGAGAGTGGCCAG 2520
Db 2421 ACAGAGAGGGGTACAGACAATGATGAGTGGACAGCCCAAAATAGAGAGTGGCCAG 2480
QY 2521 ATGGGGCTGGAGAGCGCGCCACACTGGAGTATAGAGCATCAGGACATCTCAGC 2580
Db 2481 ATGGGGCTGGAGAGCGCGCCACACTGGAGTATAGAGCATCAGGACATCTCAGC 2540
QY 2581 AGCAAGATGCCCAACCATGGGTGAACCTTGTGGAGAACCTGGACAGCCAGCCCAAA 2640
Db 2541 AGCAAGATGCCCAACCATGGGTGAACCTTGTGGAGAACCTGGACAGCCAGCCCAAA 2600
QY 2641 GTTCCAGAGCGGAGGCTCCCTGGTCCCGCGGAGGCTCCCTGCTCAGACCGGTCTA 2700
Db 2601 GTTCCAGAGCGGAGGCTCCCTGGTCCCGCGGAGGCTCCCTGCTCAGACCGGTCTA 2660
QY 2701 AGCAAGCGGCTGGAAATCAGACACTCCCTCTCCACGGGGTGTACTATAAGAGAGTAC 2760
Db 2661 AGCAAGCGGCTGGAAATCAGACACTCCCTCTCCACGGGGTGTACTATAAGAGAGTAC 2720
QY 2761 CCCAGCACTGCTCAGAGAGGCCAGCAGGCCACCTCTCAAAAGAAACACACATAAC 2820
Db 2721 CCCAGCACTGCTCAGAGAGGCCAGCAGGCCACCTCTCAAAAGAAACACACATAAC 2780
QY 2821 TCTCCATTCCTCTCACTCTCCAGAACAGAGCTTTGGCAGGGAGACANCCGCG 2880
Db 2781 TCTCCATTCCTCTCACTCTCCAGAACAGAGCTTTGGCAGGGAGACANCCGCG 2840
QY 2881 CCGCGCCCGCAGAGGCTGACTCCATCCAGGTGCACAGCTCCCGACCTCCGCGAGGCC 2940
Db 2841 CCGCGCCCGCAGAGGCTGACTCCATCCAGGTGCACAGCTCCCGACCTCCGCGAGGCC 2900
QY 2941 GTGACTGTCTGAGCAGCCCGCAGCTCAACGCTACAACTCACTGACAAATGTGGGGTG 3000
Db 2901 GTGACTGTCTGAGCAGCCCGCAGCTCAACGCTACAACTCACTGACAAATGTGGGGTG 2960
QY 3001 AAGCGTACGCGCTCGCTAAGCGCGGACGTACGCCCAACCAATCCTTTCTCCCTTCC 3060
Db 2961 AAGCGTACGCGCTCGCTAAGCGCGGACGTACGCCCAACCAATCCTTTCTCCCTTCC 3020
QY 3061 ACATCCATGAAGCCCAATCATGCTGTATATAA 3093
Db 3021 ACATCCATGAAGCCCAATCATGCTGTATATAA 3053
```

RESULT 6

AAF98469

ID AAF98469 standard; cdna; 3550 BP.

XX

AC AAF98469;

XX

DT 07-JUN-2001 (first entry)

XX

DE Human cdna clone CJ145\_1 sequence SEQ ID 160.

XX

KW Human; secreted protein; nutrient; cytokine modulator; proliferation;  
KW differentiation; immune system modulator; tissue growth; chemotactic;  
KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition; ss;

haematopoiesis.  
XX Homo sapiens.  
XX WO200119988-A1.  
XX 22-MAR-2001.  
XX 14-SEP-2000; 2000WO-US25135.  
XX 17-SEP-1999; 99US-0398829.  
XX (GSM ) GENETICS INST INC.  
XX Jacobs K, McCoy JM, LaValle ER, Collins-Racie LA, Evans C;  
XX Merberg D, Treacy M, Bowman MR, Spaulding V, Agostino MJ;  
XX WPI: 2001-244801/25.  
XX P-PSDB; AAB90731.  
XX Isolated nucleic acids encoding polypeptides, useful for modulating  
XX e.g. cytokine and cell proliferation/differentiation activity, the  
XX immune system and hematopoiesis regulating activity -  
XX Disclosure; Page 486-487; 557pp; English.  
XX Human cdna clones represented in AAF98374 - AAF98489 encode secreted  
XX proteins AAB90667 - AAB90750. The cdna clones are isolated from various  
XX tissue types, and may be used in the prevention, treatment and diagnosis  
XX of diseases associated with inappropriate protein expression. The  
XX polypeptides and nucleic acids may be used as nutrients or to modulate  
XX cytokine and cell proliferation/differentiation activity and may also be  
XX involved in modulation of the immune system. The cdna sequences,  
XX proteins, their agonists and/or antagonists exhibit haematopoiesis  
XX regulating activity; tissue growth activity; activin/inhibin activity;  
XX chemotactic/chemokinetic activity; haemostatic and thrombolytic  
XX activity; receptor/ligand activity; anti-inflammatory activity;  
XX haematopoiesis activity; cadherin/tumour suppressor activity; and/or  
XX tumour inhibition activity. Included in the invention are probes  
XX represented in AAF98490 - AAF98572 which are specific for the cdna clones  
XX encoding the secreted proteins.

Sequence 3550 BP; 957 A; 994 C; 856 G; 742 T; 1 other:

Query Match 88.88; Score 2745.2; DB 22; Length 3550;

Best local Similarity 94.58; Pred. No. 0;

Matches 2923; Conservative 1; Mismatches 4; Indels 165; Gaps 1;

Qy 1 ATGAGGTGAGAGGCTTCTGCTGCTATATTTTACACATGCTACACATTGCTGGGCTGTTTC 60  
Db 126 ATGAGGTGAGAGGCTTCTGCTGCTATATTTTACACATGCTACACATTGCTGGGCTGTTTC 165  
Qy 61 CCAGNAGATTCTGAGCCANTCAGTATTTTCCATGGCAACTATACAAACAGTATCCGGTG 120  
Db 186 CCAGNAGATTCTGAGCCANTCAGTATTTTCCATGGCAACTATACAAACAGTATCCGGTG 245  
Qy 121 TTGTGGGCCACAGCCAGGACGAGCAACACACAGAGGCGACAGGCTGGACATCCAGATG 180  
Db 246 TTGTGGGCCACAGCCAGGACGAGCAACACACAGAGGCGACAGGCTGGACATCCAGATG 305  
Qy 181 ATTATGATCATGACGGAACCCCTCTACATTTGCTCTAGGACCATATTTTACTGTTGAT 240  
Db 306 ATTATGATCATGACGGAACCCCTCTACATTTGCTCTAGGACCATATTTTACTGTTGAT 365  
Qy 241 ATAGACACATCACACACGGAAGAAATTTATTTAGCAAAAAAAGTACATGGAATCTAGA 300  
Db 366 ATAGACACATCACACACGGAAGAAATTTATTTAGCAAAAAAAGTACATGGAATCTAGA 425  
Qy 301 CAGGCCCATGTAGACACATGCAGNATGAAGGGAAAAACATGAAGTATGATCCCACTTT 360  
Db 426 CAGGCCCATGTAGACACATGCAGNATGAAGGGAAAAACATGAAGTATGATCCCACTTT 485  
Qy 361 AITAAAGTCTTCTTAAAGAAAAACGATGATGCTTTGCTCTGTGGAACATAATGCTTTC 420









[illegible]

Db	1724	GT-TGGGTCTCTACCTGTGTGTATAAAGGTTCCCTTTGGCCGGTGTGAACGACATGGGAAG	1773
Qy	1561	TGTAAATAAAACCTGTATTTCCTCCAGAGACCCCATATTGTGGATGGATAAAGGAAGTGGT	1620
Db	1774	TGTAAATAAAACCTGTATTTCCTCCAGAGACCCCATATTGTGGATGGATAAAGGAAGTGGT	1833
Qy	1621	GCTGTGAGGCATTTATCACCCACACAGACACTGACTTTTGGACGAGACATAGAGCGTGGC	1680
Db	1834	GCTGTGAGGCATTTATCA-CCCAACAGACAGACTGACTTTTGGACGAGACATAGAGCGTGGC	1893
Qy	1681	AATACAGATGGTCTGGGGACTGTCAAAATCCCTTTGTGSCACTGAAATGGCAATTCAGT	1740
Db	1894	AATACAGATGGTCTGGGGACTGTCAAAATCCCTTTGTGSCACTGAAATGGCAATTCAGT	1941
Qy	1741	TCCCTCTTGCCCAAGCACACACATCAGATTCAGAGGCTCAAGAGGGTAIGAGTCTAGG	1800
Db	1942	-	1941
Qy	1801	GGAGGAATGCTGGACTGGAAGCATCTGCTTGACTCACCTGACAGCAACAGACCCCTTTGGGG	1860
Db	1942	-	1941
Qy	1861	GGAGTGTCTTCCCATTAATCACCAAGACACAGAGGAGTGAATTCGGGAAGATTAACCTCAAA	1920
Db	1942	-	1963
Qy	1921	GGCCAGCACAGCTGGTTCGGCTACCCCTTGGCCATTCGACTGATCCTGTGGCTTTCGCTC	1980
Db	2028	GGCCAGCACAGCTGGTTCGGCTACCCCTTGGCCATTCGACTGATCCTGTGGCTTTCGCTC	2028
Qy	1981	ATGGGGCGCTCTTCCTGGGCATCAGCGTCTACTGGGTCTGTGATCATCGCGCAAGAC	2040
Db	2029	ATGGGGCGCTCTCTCTGGGCATCAGCGTCTACTGGGTCTGTGATCATCGCGCAAGAC	2088
Qy	2041	GTGGCTGTGGTGAGCGCAAGAGAGAGGAGCTCACCCACTCGCGCGGGGCTCCATGAGC	2100
Db	2089	GTGGCTGTGGTGAGCGCAAGAGAGAGGAGCTCACCCACTCGCGCGGGGCTCCATGAGC	2148
Qy	2101	AGCGTCACCAAGCTCAGCGGCCTCTTTGGGGACACTAATCCAAAGACCCAAAGCCGGAG	2160
Db	2149	AGCGTCACCAAGCTCAGCGGCCTCTTTGGGGACACTAATCCAAAGACCCAAAGCCGGAG	2208
Qy	2161	GGCATCTCAGCGCACCATGACACAGCGCAAGCTCGCCACTCCCGCAACACGGCCAG	2220
Db	2209	GGCATCTCAGCGCACCATGACACAGCGCAAGCTCGCCACTCCCGCAACACGGCCAG	2268
Qy	2221	ATGCTATTAAAGCAGACCAAGCACCACTGGAGCTGACGGCCCTCCGCCACCCAGAGTCA	2280
Db	2269	ATGCTATTAAAGCAGACCAAGCACCACTGGAGCTGACGGCCCTCCGCCACCCAGAGTCA	2328
Qy	2281	ACCCACAGCTTGAGACAGACGGAGACCCAGCGGGGACGCCGCTGAGTGGGAGAGGAAC	2340
Db	2329	ACCCACAGCTTGAGACAGACGGAGACCCAGCGGGGACGCCGCTGAGTGGGAGAGGAAC	2388
Qy	2341	CAGAACTCATCAATGGCTGCACAAAGGACATGCCCCCATGGGCTCCCTGTGATTCCTCC	2400
Db	2389	CAGAACTCATCAATGGCTGCACAAAGGACATGCCCCCATGGGCTCCCTGTGATTCCTCC	2448
Qy	2401	ACGAGCTGGCCCTTGCGGGCTCCGCCAGGCACATGCCCGAGGCTGGTGGTCTCTGCCCATC	2460
Db	2449	ACGAGCTGGCCCTTGCGGGCTCCGCCAGGCACATGCCCGAGGCTGGTGGTCTCTGCCCATC	2508
Qy	2461	ACGAGCAGGGCTACGAGCATGAGTACGTGGACACGCCCAAAATGACGAGGTGGGCCAG	2520
Db	2509	ACGAGCAGGGCTACGAGCATGAGTACGTGGACACGCCCAAAATGACGAGGTGGGCCAG	2568
Qy	2521	ATGGCGCTGGAGGACCGAGCGGCACACTGGAGTATAAGACCATCAAGGAACATCTCAGC	2580
Db	2569	ATGGCGCTGGAGGACCGAGCGGCACACTGGAGTATAAGACCATCAAGGAACATCTCAGC	2628
Qy	2581	AGCAAGAGTCCCAACCATGGGTGAACCTTGTGGAGAACCTGAGACAGCTGCCCCCCCCAAA	2640
Db	2629	AGCAAGAGTCCCAACCATGGGTGAACCTTGTGGAGAACCTGAGACAGCTGCCCCCCCCAAA	2688



Db	1059	GAGCACCAGCTCATGGATGAGGACATGGCCCTCCATCTCTTCAACAGCCCAATGGTTCCTCGAG	1112
QY	1239	AACAATGGTCAGATACCGGCTTACCAAAATTCAGTGGACACAGAGTGGTGGGCCATATCA	1295
Db	1119	AACAATGGTCAGATACCGGCTTACCAAAATTCAGTGGACACAGAGTGGGCCAATATCA	1178
QY	1299	GAATCACTGTGGTTTTCCTGAGGATCAGAGAGGGAATCATCTTGAAGTTTTCGGCCAG	1358
Db	1179	GAATCACTGTGGTTTTCCTGAGGATCAGAGAGGGAATCATCTTGAAGTTTTCGGCCAG	1238
QY	1359	AATAGGAATAGTGGTTTCTAAATGACAGCCCTTCCCTGAGGAGATCAGTGTATTAA	1418
Db	1239	AATAGGAATAGTGGTTTCTAAATGACAGCCCTTCCCTGAGGAGATCAGTGTATTAA	1298
QY	1419	CTCTGAAAATGACAGTATGATGGAGTCCGAGACAAAGGATCATGGGCATCGACGTGGA	1478
Db	1299	CTCTGAAAATGACAGTATGATGGAGTCCGAGACAAAGGATCATGGGCATCGACGTGGA	1358
QY	1479	CAGAGCAAGCAGCTCTCTGTATGCTGGCTTCCTAGCTGTGTGATAAATGTTCCGCTTGG	1538
Db	1359	CAGAGCAAGCAGCTCTCTGTATGCTGGCTTCCTAGCTGTGTGATAAATGTTCCGCTTGG	1418
QY	1539	CCGGTGTGACGACATGGGAATGTAAATAAAGCTGTATTCCTCCAGAGACCATATG	1598
Db	1419	CCGGTGTGACGACATGGGAATGTAAATAAAGCTGTATTCCTCCAGAGACCATATG	1478
QY	1599	TGGATGGATAAGAGAGGTGGTCCCTGCAGCCATTATCACCACACAGAG	1653
Db	1479	TGGATGGATAAGAGAGGTGGTCCCTGCAGCCATTATCACCACACAGAGACAGAG	1538
QY	1654	ACTTTTGAGCAGGACATAGAGGTGGACATACAGATGGTGTGGGGAGCTGTCAATTC	1713
Db	1539	AGGCTTTTGTCTTCTGTGTACCAGGCTCCACTTTACACAGATGCTATTAATCACTTC	1598
QY	1714	TTTGTGCACATGAATGGGCATTCAGATTCCCTCTTGGCCAGCACACCATCAGATTGG	1773
Db	1599	ACTATGAGACAGACATCCATCATATCACAGCAGTCAGAACGGCGCTCCAGTGA	1658
QY	1774	ACGGCTCAAGAGGGGTATGAG	1794
Db	1659	AGGGGTGAGGGGAAAGTGACACATTAACCATTCAGATAGTGTGGTTGCCCTGT	1718
QY	1795	-----TCTAGGGAGGANTGCTGGATCGAAGCATCTGTTCA	1832
Db	1719	GTCTTACAGCTTAGGGTGAATGGGGTGGGGTGTCTAAAGGCAACACATTTCCG	1778
QY	1833	CTCACCTCAGACACAGCCCTTGGGGGAGTGTCTTCCCATATCACCAACACAGAA	1892
Db	1779	CAGACGGAGCTCGGGTGGCCACAGAGGGGTGTGGGAAGATCCCACTGGAAGCCCA	1838
QY	1893	GGGAGTGATTCGGGAAAGTTATCTCAAAAGCCACGACAGAGTGGTCCGCTCACCCCTTT	1952
Db	1839	CAGAGTGAATTCGGGAAAGTTATCTCAAAAGCCACGACAGTGGTCCGCTCACCCCTTT	1898
QY	1953	GGCCATTGCACTCATCTGGCTTCGTGATGGGGGGCTCTCTCGGGCATCACCGTCA	2012
Db	1899	GGCCATTGCACTCATCTGGCTTCGTGATGGGGGGCTCTCTCGGGCATCACCGTCA	1958
QY	2013	CTGGCTGTGATCATCGGCCAAAGAGCGTGGCTGTCAGCGCAAGGAGAGAGGT	2072
Db	1959	CTGGCTGTGATCATCGGCCAAAGAGCGTGGCTGTCAGCGCAAGGAGAGAGGT	2018
QY	2073	CACCCACTCGCGCGGGGTTCTATGACAGAGCTGACATAGCTTAGCGGCTCTTTGGGGA	2132
Db	2019	CACCCACTCGCGCGGGGTTCTATGACAGAGCTGACATAGCTTAGCGGCTCTTTGGGGA	2078
QY	2133	CACCTCAATCCGAAGACCCAAAGCGGGAGGCGCATGCTCAGCGCACTCATGCAACAGCCAA	2192
Db	2079	CACCTCAATCCGAAGACCCAAAGCGGGAGGCGCATGCTCAGCGCACTCATGCAACAGCCAA	2138
QY	2193	GCTCGCCACTCCCGGCACACGGCCAAAGTGTCTATTAAGACAGACAGACCAACCTGGA	2252
Db	2139	GCTCGCACTCCCGGCACACGGCCAAAGTGTCTATTAAGACAGACAGACCAACCTGGA	2198

Qy	2253	CCTGACGGGCGCTCCGACCCGAGTGAACCCGCAACGCTCGACGAGGCGGAGCCGAG	2312
Db	2199	CCTGACGGGCGCTCCGACCCGAGTGAACCCGCAACGCTCGACGAGGCGGAGCCGAG	2258
Qy	2313	CCGCGGCACCGCGAGTGGGAGGAGNCCAGAACCTCATCAATCCCTGCACNAGGACAT	2372
Db	2259	CCGCGGCAGCGCGAGTGGGAGGAGAACCAAGAACCTCATCAATGCCCTGCACNAGGACAT	2318
Qy	2373	GCGCCCATGGGCTCCGCTGTGATTCGACAGAGACCTCCGCTCGGGCCCTCCGCGACCA	2432
Db	2319	GCGCCCATGGGCTCCGCTGTGATTCGACAGGACCTCCGCTCGGGGCCCTCCGCGACCA	2378
Qy	2433	CATCCGACGGTGGTGGTCTGCTGCCCATACGACAGCAGGGCTACACAGCATGAGTACGTGA	2492
Db	2379	CATCCGACGGTGGTGGTCTGCTGCCCATACGACAGCAGGGCTACACAGCATGAGTACGTGA	2438
Qy	2493	CCAGCCCAAAATGAGCGAGTGGGCCAGATGGCGCTGAGAGACACAGGCGCCGACACTGGA	2552
Db	2439	CCAGCCCAAAATGAGCGAGTGGGCCAGATGGCGCTGAGAGACACAGGCGCCGACACTGGA	2498
Qy	2553	GTATAGACCATCAGGAACATCTCAGACGACAGCTCCCAACCATGGSGTGACCTTGT	2612
Db	2499	GTATAGACCATCAAGGAACATCTCAGACGACAGGATCCCAACCATGGSGTGAACCTTGT	2558
Qy	2613	GGAGACCTCCCTGTCTCAGACCGGTCTAAGCAAGCGGTGGAAATGCACCACTCTCTTC	2672
Db	2559	GGAGACCTCCCTGTCTCAGACCGGTCTAAGCAAGCGGTGGAAATGCACCACTCTCTTC	2618
Qy	2673	GGAGCGTCCCTGTCTCAGACCGGTCTAAGCAAGCGGTGGAAATGCACCACTCTCTTC	2732
Db	2619	GGAGCGTCCCTGTCTCAGACCGGTCTAAGCAAGCGGTGGAAATGCACCACTCTCTTC	2678
Qy	2733	CTACGGGGTTGACTATAGAGGAGCTACCCGACGAACTGCGCTCAGGAGAACCCAGGC	2792
Db	2679	CTACGGGGTTGACTATAGAGGAGCTACCCGACGAACTGCGCTCAGGAGAACCCAGGC	2738
Qy	2793	CACCACCTCAAAAAGAACACACTAACTCTCCAAATTCCTCTCACCTCTCCAGAAACA	2852
Db	2739	CACCACCTCAAAAAGAACACACTAACTCTCCAAATTCCTCTCACCTCTCCAGAAACA	2798
Qy	2853	GAGCTTTGGCAGGGGAGAACACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	2912
Db	2799	GAGCTTTGGCAGGGGAGAACACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	2856
Qy	2913	GCACAGTCCGAGGCATCTGGCAGGCGGTGACTGTCTCGAGGAGAGCCGACCCCTCAAGC	2972
Db	2859	GCACAGTCCGAGGCATCTGGCAGGCGGTGACTGTCTCGAGGAGAGCCGACCCCTCAAGC	2918
Qy	2973	CTACAACCTCACTGCACAAGTCCGGGGCTGAAGCGTACGCGCTCGGTAAAGCGCGAGCTACC	3038
Db	2919	CTACAACCTCACTGCACAAGTCCGGGGCTGAAGCGTACGCGCTCGGTAAAGCGCGAGCTACC	2972
Qy	3033	CCCCAAGCATCTCTTGTCTCCCTTTCACATTCGCAATGAAGGCCCAATGATGGGTGTACATA	3092
Db	2979	CCCCAAGCATCTCTTGTCTCCCTTTCACATTCGCAATGAAGGCCCAATGATGGGTGTACATA	3038
Qy	3033	A 3093	
Db	3039	A 3039	
RESULT 9			
AAS89721			
CD	AAS89721 standard; cDNA: 3039 BP.		
XX	AAS89721;		
XX			
DT	13 FEB-2002 (first entry)		
XX	DNA encoding novel human diagnostic protein #25525.		
DE	Human: chromosome mapping; gene mapping; gene therapy; forensic;		
XX			
KW	Human: chromosome mapping; gene mapping; gene therapy; forensic;		

DNA encoding novel human diagnostic protein #25525.  
Human: chromosome mapping; gene mapping; gene therapy; forensic





CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 XX of the present invention.

SQ Sequence 3041 BP; 804 A; 886 C; 714 G; 637 T; 0 other;

Query Match 70.6%; Score 2182.4; DB 22; Length 3041;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 910 CGTATCAACGGCGGTGATGTTGCTGCGCAAGG-TTTCTACACGTTATAACACATCCCT 969  
 DB 1 CGTATCAACGGCGAIGATGTTGCTGCGCAAGGTTTCTACACGTTATAACACATCCCT 60  
 QY 970 GGGTCTGACATCTGTGCTATGACATCTTGACATTCACATGCTTTTACTGGGAGATTC 1029  
 DB 61 GGGTCTGACATCTGTGCTATGACATCTTGACATTCACATGCTTTTACTGGGAGATTC 120  
 QY 1030 AAGCAACAGAGTCTCTGATTCACCTGAGACACACAGTCTCTGATGACAGTTCCTTAAG 1069  
 DB 122 AAGCAACAGAGTCTCTGATTCACCTGAGACACACAGTCTCTGATGACAGTTCCTTAAG 180  
 QY 1090 CCCAGGCGAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149  
 DB 181 CCCAGGCGAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 QY 1150 TCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1209  
 DB 241 TCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 QY 1210 TCCATCTTCAACAGGCGATGTTCTGAGAGAGATGCTGAGATACCGCTTACCAAAAT 1269  
 DB 301 TCCATCTTCAACAGGCGATGTTCTGAGAGAGATGCTGAGATACCGCTTACCAAAAT 360  
 QY 1270 GCAGTGGACACAGTGTGTGGGCAATATCAATACACACTGTGGTTTCTGGGATCAGAG 1329  
 DB 361 GCAGTGGACACAGTGTGTGGGCAATATCAATACACACTGTGGTTTCTGGGATCAGAG 420  
 QY 1330 AAGGGAATCATCTGATGTTTGGCCAGATAGAGAAATAGTGTGTTTCTAAATGACAGC 1389  
 DB 421 AAGGGAATCATCTGATGTTTGGCCAGATAGAGAAATAGTGTGTTTCTAAATGACAGC 480  
 QY 1390 CTTTCTGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1449  
 DB 481 CTTTCTGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
 QY 1450 GACAAAAGGATCATGGGCA-TGACGCTGGACAGACCAAGCAGCTCTGTGTGTGGCTTC 1509  
 DB 541 GACAAAAGGATCATGGGCA-TGACGCTGGACAGACCAAGCAGCTCTGTGTGTGGCTTC 600  
 QY 1510 TCTACCTGTGTGATAAGGTT-CCCTTGGCGGTGTGAAACACATGCGAGTGTAAATAA 1569  
 DB 601 TCTACCTGTGTGATAAGGTT-CCCTTGGCGGTGTGAAACACATGCGAGTGTAAATAA 660  
 QY 1570 ACCTGTATTCCTCCAGAGACCCATATGTGGATGATTAAGGAAGTGTGCTGCGAGC 1629  
 DB 661 ACCTGTATTCCTCCAGAGACCCATATGTGGATGATTAAGGAAGTGTGCTGCGAGC 720  
 QY 1630 CATTTATCAACCAACAGCAGACTGACITTTGACAGGACATAGAGCGTGGCAATACAGAT 1689  
 DB 721 CATTTATCAACCAACAGCAGACTGACITTTGACAGGACATAGAGCGTGGCAATACAGAT 750  
 QY 1690 GGTCGGGGAGTCTGACAAATTCCTTTGGCACTGAATGGGATTCAGATTCCTTCCTTC 1749  
 DB 781 GGTCGGGGAGTCTGACAAATTCCTTTGGCACTGAATGGGATTCAGATTCCTTCCTTC 840  
 QY 1750 CCCAGCACACCAATCAGATTCGACGGCTCAAGAGGGGATGATAGTCTAGGGAGGATTC 1809  
 DB 841 CCCAGCACACCAATCAGATTCGACGGCTCAAGAGGGGATGATAGTCTAGGGAGGATTC 900

QY 1810 CTGGACTGGAAGCATCTGCTTACTCACCIGACAGACAGACCCCTTTGGGGGAGTGCT 1869  
 DB 901 CTGGACTGGAAGCATCTGCTTACTCACCIGACAGACAGACCCCTTTGGGGGAGTGCT 960  
 QY 1870 TCCATATAA-TACCAACAGAGAGGAGTATTTCGGAAAG-TACTTCAAGAGCCACAGAC 1929  
 DB 961 TCCATATAA-TACCAACAGAGAGAGGAGTATTTCGGAAAG-TACTTCAAGAGCCACAGAC 1020  
 QY 1930 CAGCTGGTTCCCGTCAACCTCTTTGGCCATTGCGACTTCTGGCTTCTGCTAIGGGGGC 1989  
 DB 1021 CAGCTGGTTCCCGTCAACCTCTTTGGCCATTGCGACTTCTGGCTTCTGCTAIGGGGGC 1080  
 QY 1990 GTCTTCTGGGATCACCGTCTACTCGCTCTGTGATCATCGGCGCAAGACAGTGGCTG 2049  
 DB 1081 GTCTTCTGGGATCACCGTCTACTCGCTCTGTGATCATCGGCGCAAGACAGTGGCTG 1140  
 QY 2050 GTGACGCGAAGAGAGAGGAGCTCACCCACTCGCGCGGGCTTCCATGAGCAGCTCAC 2109  
 DB 1141 GTGACGCGAAGAGAGAGGAGCTCACCCACTCGCGCGGGCTTCCATGAGCAGCTCAC 1200  
 QY 2110 AAGCTCAGGGGCTCTTTGGGACACTCAATCCAAAGACCCAAAGCCGAGGCGCATCTC 2169  
 DB 1202 AAGCTCAGGGGCTCTTTGGGACACTCAATCCAAAGACCCAAAGCCGAGGCGCATCTC 1260  
 QY 2170 AGCCCACTCATGCACAACCGCAAGCTTCGCCACTTCGCCGCAACACGCGCAAGATGCTCAT 2229  
 DB 1261 AGCCCACTCATGCACAACCGCAAGCTTCGCCACTTCGCCGCAACACGCGCAAGATGCTCAT 1320  
 QY 2230 AAAGCAGACAGCACCACCTGACCTGAGCGGCTTCCACCCAGAGTCAACCCCAAG 2289  
 DB 1321 AAAGCAGACAGCACCACCTGACCTGAGCGGCTTCCACCCAGAGTCAACCCCAAG 1380  
 QY 2290 GTGACGACAGAGGAGAGCGCGCGCGAGCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAG 2349  
 DB 1382 GTGACGACAGAGGAGAGCGCGCGCGAGCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAG 1440  
 QY 2350 ATCAATCCCTGCACAAAGGACAT-GCCCGCCCATGGGTCCCTCTGTGATTCACCGAGCCTG 2409  
 DB 1441 ATCAATCCCTGCACAAAGGACATGCCCCCATGGGTCCCTCTGTGATTCACCGAGCCTG 1500  
 QY 2410 CCCTCGGGGCTTCGCCAGGACATCCCGAGCTGTGGTCTCTGCTGCTGCTGCTGCTGCTGCT 2469  
 DB 1501 CCCTCGGGGCTTCGCCAGGACATCCCGAGCTGTGGTCTCTGCTGCTGCTGCTGCTGCTGCT 1560  
 QY 2470 GGCTACACAGTACGATGAGTGGACAGCCCAAAATGAGCGAGTGGCCAGATGGCGT 2529  
 DB 1561 GGCTACACAGTACGATGAGTGGACAGCCCAAAATGAGCGAGTGGCCAGATGGCGT 1620  
 QY 2530 GAGGACAGGCGGCCACACTGGAGTATAAGACCATCAAGGACATCTCAGCAGCAGAGT 2589  
 DB 1621 GAGGACAGGCGGCCACACTGGAGTATAAGACCATCAAGGACATCTCAGCAGCAGAGT 1680  
 QY 2590 CCAACCATGGGTGAACTTGTGGAGACCTTGGACAGCTGCCCGCCCAAGATTCCACAG 2649  
 DB 1681 CCAACCATGGGTGAACTTGTGGAGAACCTTGTGGAGAACCTTGGACAGCTGCCCGCCCAAGATT 1740  
 QY 2650 CGGAGGCGCTCCCTGGGTCCCGCGGAGCTTCCCTGCTCAGACGGTCTAGCAGAGCG 2709  
 DB 1741 CGGAGGCGCTCCCTGGGTCCCGCGGAGCTTCCCTGCTCAGACCGGTCTAGCAGAGCG 1800  
 QY 2710 CTGGAATGCACACTCTCTTCTACGGGGTGTGACTATAAGAGGAGCTACCCCGAGAC 2769  
 DB 1801 CTGGAATGCACACTCTCTTCTACGGGGTGTGACTATAAGAGGAGCTACCCCGAGAC 1860  
 QY 2770 TGCTCAGAGAGCCACAGGCGCCACTCTCAAAAGAAACAACTAACTCTCTCCAT 2829  
 DB 1861 TGCTCAGAGAGCCACAGGCGCCACTCTCAAAAGAAACAACTAACTCTCTCCAT 1920  
 QY 2830 TCTCTCAGCTCTCCAGAAACAGAGCTTTGGCAGGGAGACAAACCGCGCGCGCGCGCG 2889  
 DB 1921 TCTCTCAGCTCTCCAGAAACAGAGCTTTGGCAGGGAGACAAACCGCGCGCGCGCGCG 1980  
 QY 2890 CAGAGGGTGGACTCCATCCAGGTGCACAGCTCCAGCCACTCTGGCAGGCCCTGACTGTC 2949

```

Db 1981 CAGAGGTGACCTCCATCCAGGTGACAGCTCCAGCCATCTGCGCAGSCCGTACCTGTC 2040
Qy 2950 TCGAGGAGCCAGCCCTCAAGGCTACAACTCACTGACAAAGTGGGGTGAAGCGTAGS 3009
Db 2041 TCGAGGAGCCAGCCCTCAAGGCTACAACTCACTGACAAAGGTCGGGCTGAAGCGTAGS 2100
Qy 3010 CCTCGCTAAGCGGAGCTACCCGCGCAACAGCATCTTGGCTGCCCTTCCACATGCCATG 3069
Db 2101 CCTCGCTAAGCGGAGCTACCCGCGCAACAGCATCTTGGCTGCCCTTCCACATGCCATG 2160
Qy 3070 AAGCCCAATGATGGTGATACATA 3093
Db 2161 AAGCCCAATGATGGTGATACATA 2184

RESULT 11
ID AAA93630 standard: cDNA; 1840 BP.
XX
AC AAA93630;
XX
DT 16-JAN-2001 (first entry)
DE Human scaphorin protein-like splice variant SEX_ pCR2.1-2864933 cDNA.
XX
KW SEX protein; human; secreted; membrane-associated; cancer;
KW proliferation regulator; differentiation regulator; non-malignant tumour;
KW immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;
KW infection; inflammatory disorder; arthritis; haematopoietic disorder;
KW skin disorder; cardiovascular disorder; atherosclerosis; restenosis;
KW neurological disease; Alzheimer's disease; trauma; wounding;
KW spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;
KW anti-HIV; anti-inflammatory; antiarthritic; antiarteriosclerotic;
KW neuroprotective; vulnerary; antiallergic; antimicrobial; cardiant;
KW dermatological; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200053742-A2.
XX
PD 14-SEP-2000.
XX
PP 09-MAR-2000; 2000WO-US06280.
XX
PR 09-MAR-1999; 99US-0123667.
XX
PR 08-MAR-2000; 2000US-0123667.
XX
PA (CURA-) CURAGEN CORP.
XX
P1 Shimkets RA;
XX
WP1: 2000-5943:8/56.
XX
P-PSDB: AAB23043.
XX
PT Novel human membrane associated or secreted polypeptides and
PT polynucleotides useful for diagnosis, prevention and treatment of
PT pathological states such as cancer, immune, cardiovascular and
PT neurological disorders.
XX
PS Claim 3; Fig 15; 151pp; Eng.ish.
XX
CC Sequences AAA93616-A93631 and AAA93673-A93676 represent nucleic acids
CC which encode human SEX proteins (AAB23029-B23045). The SEX proteins
CC of the invention are either secreted or membrane-associated proteins
CC and act as regulator of cellular proliferation and differentiation. SEX
CC proteins or nucleotides are useful for diagnosing the presence of, or
CC predisposition to, a disease associated with altered levels of SEX
CC proteins and nucleotides. The SEX proteins are also useful to screen
CC compounds that modulate SEX activity or expression. The interaction of
CC a SEX protein with other cellular proteins may be useful to modulate
CC the activity of a partner protein, cellular proliferation, cellular
CC differentiation and cell survival. SEX nucleotides are useful for the

```

```

CC recombinant expression of SEX protein, and may be used to detect SEX mRNA
CC or genetic lesions in the SEX gene. They may also be used to modulate
CC SEX expression (e.g., using antisense oligonucleotides). SEX nucleic
CC acid sequences are also useful for identifying a cell or tissue type in
CC a biological sample, and in forensic biology. SEX primers or probes are
CC useful for detecting the presence of SEX nucleotides and for screening
CC tissue cultures for contamination. Diseases that may be treated or
CC prevented using SEX proteins or nucleotides include cancer (e.g.,
CC colorectal carcinoma, prostate cancer), benign tumours, immune disorders
CC (including autoimmune diseases, transplant rejection, allergies, AIDS),
CC infections, inflammatory disorders, arthritis, haematopoietic disorders,
CC skin disorders, cardiovascular disorders, atherosclerosis, restenosis,
CC neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,
CC surgical or traumatic wounds, spinal cord injury), and skeletal
CC disorders.
XX
SQ Sequence 1890 BP; 535 A; 426 C; 463 G; 466 T; 0 other;

```

```

Query Match: 60.8%; Score 1841.6; DB 21; Length 1890;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1884; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy 49 GGGGCTGGTTTCCGAGAGATTCTCAGCCAATCAGTAATTTCGATGCGCACTATACAAA 108
Db 1 GGA:CCG:TTCCGAGAAGATTCAGCCAATCAG:ATTTCCGATGCGCACTATACAAA 60
Qy 109 CAGTATCCGGTGTTTGGGCCCAAGCCAGGAGCGGAACCCACACAGAGGCGACAGGCTG 168
Db 61 CAGTATCCGGTGTTTGGGCCCAAGCCAGGAGCGGAACCCACACAGAGGCGACAGGCTG 120
Qy 169 GACATCCAGATGATATGATCATCAAGCGGAACCTCTACAT:TCGTGCTAGGACCATATT 228
Db 121 GACATCCAGATGATATGATCATCAAGCGGAACCTCTACAT:TCGTGCTAGGACCATATT 180
Qy 229 TATACITGTGATATAGACACATCACACAGGAGAAATTTATTCTAGCAAAAACACTGACA 288
Db 131 TATACITGTGATATAGACACATCACACAGGAGAAATTTATTGTACAAAACACTGACA 240
Qy 289 TGGAAATCTAGACAGGCGGATGTAGACACATGCAAGATGAAGGAGAAACATAGGATGAG 348
Db 241 TGGAAATCTAGACAGGCGGATGTAGACACATGCAAGATGAAGGAGAAACATAGGATGAG 300
Qy 349 TGGCAACACTTTTAAAGTTCTTCTAAAGAAAACGATGATGCATTTGTTGCTGTGGA 408
Db 301 TCCCAACAACCTTTTAAAGTTCTTCTAAAGAAAACGATGATGCATTTGTTGCTGTGGA 360
Qy 409 ACTAATGCCCTTCAACCCCTTCCTGCAGAACTATAGATGGATACATTTGGAGCATTGCG 468
Db 361 ACTAATGCCCTTCAACCCCTTCCTGCAGAACTATAGATGGATACATTTGGAGCATTGCGG 420
Qy 469 GATGAATTCAGCGGAATGGCCAGATGCCCATATGATGCCAAACATGCCAACGTTGCACTG 528
Db 421 GATGAATTCAGCGGAATGGCCAGATGCCCATATGATGCCAAACATGCCAACGTTGCACTG 480
Qy 529 TTTCAGATGGAAAACATATACTCAGCCACAGTGCATCTACTTCCTTGCCTTGCAGGAGTC 588
Db 481 TTTCAGATGGAAAACATATACTCAGCCACAGTGCATCTACTTCCTTGCCTTGCAGGAGTC 540
Qy 589 ATTACCGGAGCTCTTGGAGAAAGCCCTACCCCTGCGGAGCCGCTCAAGCAGCATTCAAAATGG 648
Db 541 ATTACCGGAGCTCTTGGAGAAAGCCCTACCCCTGCGGAGCCGCTCAAGCAGCATTCAAAATGG 600
Qy 649 TTGAAGAACCACTACTTGTTCAGGCGGTGGATTTACGAGATATATATCTACTTCTTCTTC 708
Db 651 TTGAAGAACCACTACTTGTTCAGGCGGTGGATTTACGAGATATATATCTACTTCTTCTTC 660
Qy 709 AGGGAATAGCAGTGGACTATAACACCATGGGAAGTAGTTTTTCCCAAGATGGCTCAG 768
Db 661 AGGGAATAGCAGTGGAGTATAACACCATGGGAAGTAGTTTTTCCCAAGATGGCTCAG 720
Qy 769 GTTTGTAGAAGATGATGGGAGGATCTCAAGAGTCTCTGGAGAAACAGTGGAGCTCTTC 828
Db 721 GTTTGTAGAAGATGATGGGAGGATCTCAAGAGTCTCTGGAGAAACAGTGGAGCTCTTC 780

```



QY 829 CTGAAGGCGCGTTGAAGTGTCTAGTTCCTGGAGACTCTCAATTTTATTCAACATTC 888  
DB |||||||  
QY 781 CTGAAGGCGCGTTGAAGTGTCTAGTTCCTGGAGACTCTCAATTTTATTCAACATTC 840  
DB |||||||  
QY 889 CAGGAGATACAGATGTAATGGTATCAACAGGCGGTGAAGTGTGCTGGCAAGCTTCT 948  
DB |||||||  
QY 841 CAGGAGATACAGATGTAATGGTATCAACAGGCGGTGAAGTGTGCTGGCAAGCTTCT 900  
DB |||||||  
QY 949 ACACCTTATAACAGCAATCCCTGGGCTGCAGTCTGTGCTATGACATGCTTGACATTC 1005  
DB |||||||  
QY 901 ACACCTTATAACAGCAATCCCTGGGCTGCAGTCTGTGCTATGACATGCTTGACATTC 960  
DB |||||||  
QY 1009 AGTGTTTACTGGAGATTCAGGAACAGAGAGTCTCTGATTCGACCTGAGACAGATT 1068  
DB |||||||  
QY 961 AGTGTTTACTGGAGATTCAGGAACAGAGAGTCTCTGATTCGACCTGAGACAGATT 1020  
DB |||||||  
QY 1069 CCTGATGAACAGAGTCTCTAAGCCAGGCGAGTGTGCTGCTGAGCTATCTCTCTAGAA 1128  
DB |||||||  
QY 1021 CCTGATGAACAGAGTCTCTAAGCCAGGCGAGTGTGCTGCTGAGCTATCTCTCTAGAA 1080  
DB |||||||  
QY 1129 AGATATGCAACCTCCAAATGATTCCTGATGATACCTGAACTTCATCAAGACGACCG 1188  
DB |||||||  
QY 1081 AGATATGCAACCTCCAAATGATTCCTGATGATACCTGAACTTCATCAAGACGACCG 1140  
DB |||||||  
QY 1189 GTATGGATGAGGCAATGCTTCCATCTTCACAGGCGATGTTCTGAGAACATGGTC 1248  
DB |||||||  
QY 1241 CTGATGGATGAGGCAATGCTTCCATCTTCACAGGCGATGTTCTGAGAACATGGTC 1200  
DB |||||||  
QY 1249 AGATAGCGGCTTACCAAAATTCAGTGTGACAGCACTGTCTGGGCGATATCAGAAATC 1308  
DB |||||||  
QY 1201 AGATAGCGGCTTACCAAAATTCAGTGTGACAGCACTGTCTGGGCGATATCAGAAATC 1260  
DB |||||||  
QY 1309 GTGGTTTCTGGGATCAGCAAGGAGATCATCTGAAAGTTTGGGCGAGATAGGAAT 1368  
DB |||||||  
QY 1261 GTGGTTTCTGGGATCAGCAAGGAGATCATCTGAAAGTTTGGGCGAGATAGGAAT 1320  
DB |||||||  
QY 1369 AGTGGTTTCTAAATGACAGCGCTTTTCTGAGGAGATGAGTGTATCAACTCTGMAAA 1428  
DB |||||||  
QY 1321 AGTGGTTTCTAAATGACAGCGCTTTTCTGAGGAGATGAGTGTATCAACTCTGMAAA 1380  
DB |||||||  
QY 1429 TCACGATGATGAGTGTGAAGACAAAAGGATCATGGGCGATGAGCTGGACAGAGCAAGC 1488  
DB |||||||  
QY 1381 TCACGATGATGAGTGTGAAGACAAAAGGATCATGGGCGATGAGCTGGACAGAGCAAGC 1440  
DB |||||||  
QY 1489 AGCTCTCTATGTGGCTGTCTACCTGTGTGATAAAGTTTCCCTTGGCGGTGTGAA 1548  
DB |||||||  
QY 1441 AGCTCTCTATGTGGCTGTCTACCTGTGTGATAAAGTTTCCCTTGGCGGTGTGAA 1500  
DB |||||||  
QY 1549 CGACATGGGAAGTGTAAAAAACCTGTATTGCTCCAGAGACCCATATGTGGATGGATA 1608  
DB |||||||  
QY 1501 CGACATGGGAAGTGTAAAAAACCTGTATTGCTCCAGAGACCCATATGTGGATGGATA 1560  
DB |||||||  
QY 1609 AAGGAAGTGTGGCTGGAGTCATTTATCACTCAACAGCAGCTGACTTTTGGACGAGAC 1668  
DB |||||||  
QY 1561 AAGGAAGTGTGGCTGGAGTCATTTATCACTCAACAGCAGCTGACTTTTGGACGAGAC 1620  
DB |||||||  
QY 1669 ATAGAGCTGGCAATACAGATGCTGCGGGGACTGTCAAAATTCCTTTGTGGCACTGAAT 1728  
DB |||||||  
QY 1621 ATAGAGCTGGCAATACAGATGCTGCGGGGACTGTCAAAATTCCTTTGTGGCACTGAAT 1680  
DB |||||||  
QY 1729 GGGCATTCCAGTTCCTCTTSCCAGCAGCAACCATCATGAGTTGACGGCTCAGAGCGG 1788  
DB |||||||  
QY 1681 GGGCATTCCAGTTCCTCTTSCCAGCAGCAACCATCATGAGTTGACGGCTCAGAGCGG 1740  
DB |||||||  
QY 1789 TATGACTCTAGGGGAGGAATGCTGGAATGGAAGCATCTGCTGTGACTCACCTGACAGCACA 1848  
DB |||||||  
QY 1741 TATGACTCTAGGGGAGGAATGCTGGAATGGAAGCATCTGCTGTGACTCACCTGACAGCACA 1800  
DB |||||||  
QY 1849 GACCCCTTGGGGCGAGTGTCTTCCCATATATCAACAGCAGCAAGAGGGAGTATTCGGCAA 1908  
DB |||||||  
QY 1801 GACCCCTTGGGGCGAGTGTCTTCCCATATATCAACAGCAGCAAGAGGGAGTATTCGGCAA 1860  
DB |||||||

QY 1909 AGTTACTCAAGAGGCTACAGCAGCAGCTGG 1936  
DB |||||||  
QY 1861 AGTTACTCAAGAGGCTACAGCAGCAGCTGG 1888  
DB |||||||

## RESULT 12

AAH15834  
ID AAH15834 standard; cDNA; 2293 BP.

XX AAH15834;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:14327.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EF1074617-A2.

XX C7-FEB-2001.

XX 28-JUL-2000; 2000FP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-030253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Qia T. Isogai T. Nishikawa T. Hayashi K. Saito K. Yamamoto J;

XX Saito S. Sugiyama T. Wakamatsu A. Nagai K. Otsuki T;

XX WPI; 2001-3:8749/34.

XX Claim 8; SEQ ID 14327; 2537pf + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

XX full-length cDNAs defined in the specification. Where a primer set

XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in

XX the specification. The primer sets can be used in antisense therapy and

XX in gene therapy. The primers are useful for synthesizing polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

XX the full-length cDNAs. The primers allow obtaining of the full-length

XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

XX represent oligonucleotides, all of which are used in the exemplification

XX of the present invention.

XX Sequence 2293 BP; 579 A; 713 C; 581 G; 420 T; 0 other;

Query Match 59.0%; Score 1824.4; DB 22; Length 2293;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1825; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1268 TTGAGTGGACACAGGCGCTGGGCCATAACAGATACAGCTGCTGTTTCTGGGATCAG 1327
DB 1 TTTGGTGGGACACAGGCTGCTGGGCCATATCAGAAATCAAGCTGGTTTCTGGGATCAG 60
QY 1328 AGAAGGGAATCACTTGAAGTTTGGCCAGANTAGAAATAGTGTTTCTAAATGACA 1387
DB 61 AGAAGGGAATCACTTGAAGTTTGGCCAGANTAGAAATAGTGTTTCTAAATGACA 120
QY 1388 GCTTTCTGGAGACATGATGATGTTTACAACTGTGAANAATGCACTATGATGAGTGG 1447
DB 121 GCTTTCTGGAGACATGATGATGTTTACAACTGTGAANAATGCACTATGATGAGTGG 180
QY 1448 AAGACAAAGGATCATGGGATGATGATGATGATGATGATGATGATGATGATGATG 1507
DB 181 AAGACAAAGGATCATGGGATGATGATGATGATGATGATGATGATGATGATGATG 240
QY 1508 TCTTACTGTTGATTAAGTTTCCCTTGGCGGTGTGAACACATGGGAATGTAAAA 1567
DB 241 TCTTACTGTTGATTAAGTTTCCCTTGGCGGTGTGAACACATGGGAATGTAAAA 300
QY 1568 AAATCTGATTTGCTCCAGACACCCATATTTGATGATGATGATGATGATGATGATG 1627
DB 301 AAATCTGATTTGCTCCAGACACCCATATTTGATGATGATGATGATGATGATGATG 360
QY 1628 GCAATTTATACCCAAACAGACAGCTGATCTTTGAGCAGGACATAGAGCGTGGCAATACAG 1687
DB 361 GCAATTTATACCCAAACAGACAGCTGATCTTTGAGCAGGACATAGAGCGTGGCAATACAG 420
QY 1688 ATGTCTGGGGATGTGTCACATTCCTTTGTGGACTGAATGGCATTCAGTTCCCTCT 1747
DB 421 ATGTCTGGGGATGTGTCACATTCCTTTGTGGACTGAATGGCATTCAGTTCCCTCT 480
QY 1748 TGCCAGACACAAACACATCACATTCAGAGCGCTCAAGAGGGGTATGAGTCTAGGGAGGAA 1807
DB 481 TGCCAGACACAAACACATCACATTCAGAGCGCTCAAGAGGGGTATGAGTCTAGGGAGGAA 540
QY 1808 TGCTGGACTGGAAGCATCTGCTTGACTACCTGACAGCAGACAGCCCTTTGGGGCAGTGT 1867
DB 541 TGCTGGACTGGAAGCATCTGCTTGACTACCTGACAGCAGACAGCCCTTTGGGGCAGTGT 600
QY 1868 CTTCCTCATATCACACAGACAAAGAGGAGTGTGCGGAAAGTTACCTCAAGAGCCACG 1927
DB 601 CTTCCTCATATCACACAGACAAAGAGGAGTGTGCGGAAAGTTACCTCAAGAGCCACG 660
QY 1928 ACCAGCTGGTTCCCGTCAACCTCTTGCCATTCAGATCATCCCTGGCTTGTGATGGGG 1987
DB 661 ACCAGCTGGTTCCCGTCAACCTCTTGCCATTCAGATCATCCCTGGCTTGTGATGGGG 720
QY 1988 CCGTCTTCGGGATCAGCGTCTACTGCTGCTGATCATCGCCGCAAGAGCTGGCTG 2047
DB 721 CCGTCTTCGGGATCAGCGTCTACTGCTGCTGATCATCGCCGCAAGAGCTGGCTG 780
QY 2048 TGGTGCAGCGCAAGGAGAGGATGACCTGCTGCGCGGAGGCTGCAATGAGCAGGCTCA 2107
DB 781 TGGTGCAGCGCAAGGAGAGGATGACCTGCTGCGCGGAGGCTGCAATGAGCAGGCTCA 840
QY 2128 CCAAGCTCAGCGGCTCTTTGGGACACTCAATCCAAAGACGCAAGCCGAGGACATCC 2167
DB 841 CCAAGCTCAGCGGCTCTTTGGGACACTCAATCCAAAGACGCAAGCCGAGGACATCC 900
QY 2168 TCACGCCACTCATGCACAAACGTCAGGTGGCCACTGCGGCAACAGCGCCAGATGCTCA 2227
DB 901 TCACGCCACTCATGCACAAACGTCAGGTGGCCACTGCGGCAACAGCGCCAGATGCTCA 960
QY 2228 TTAAGACAGACACACACCTGAGCTGTATGGCCCTCCCCACCCAGAGCTCAACCCCAA 2287
DB 961 TTAAGACAGACACACACCTGAGCTGTATGGCCCTCCCCACCCAGAGCTCAACCCCAA 1020
QY 2288 CCGTCCAGCAGAAAGCGGAGCGAGCGCGGAGCGCGGAGTGGGAGAGCAACAGAAC 2347
DB 1021 CCGTCCAGCAGAAAGCGGAGCGAGCGCGGAGCGCGGAGTGGGAGAGCAACAGAAC 1080
```

```
QY 2348 TCATCAATGCTGTCACAAAGAGCATGCCGCCCATGTCGCTCCCTGCTGTGATTCCACAGGAC 2407
DB 1081 TCATCAATGCTGTCACAAAGAGCATGCCGCCCATGTCGCTCCCTGCTGTGATTCCACAGGAC 1140
QY 2408 TGCCCTTGCGGGCTGCTCCAGCAGACATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 2467
DB 1141 TGCCCTTGCGGGCTGCTCCAGCAGACATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 2468 AGGCTACCAAGCATGATGCTGAGCAGCAGCCAAAATGAGCGAGGTGGCCAGATGGCC 2527
DB 1201 AGGCTACCAAGCATGATGCTGAGCAGCAGCCAAAATGAGCGAGGTGGCCAGATGGCC 1260
QY 2528 TGGAGGACAGCGCGCACACTGAGTATAGACCATCAGGAGCATCTCAGCAGCAACA 2587
DB 1261 TGGAGGACAGCGCGCACACTGAGTATAGACCATCAGGAGCATCTCAGCAGCAACA 1320
QY 2588 GTCCCAACCATGCGGTGAGCTGTGAGAACTGTGAGACCTGCGCCCAAGTTCCAC 2647
DB 1321 GTCCCAACCATGCGGTGAGCTGTGAGAACTGTGAGACCTGCGCCCAAGTTCCAC 1380
QY 2648 AGCGGAGAGCTGCTCGGTCCCGCGGAGCTCCCTGCTCAGACGGTCTAAGCAAGC 2707
DB 1381 AGCGGAGAGCTGCTCGGTCCCGCGGAGCTCCCTGCTCAGACGGTCTAAGCAAGC 1440
QY 2708 GGCTGGAATGACACACTGCTCTTCTTACGGGTGCTGACTATAAGAGGAGCTACCCACCA 2767
DB 1441 GGCTGGAATGACACACTGCTCTTCTTACGGGTGCTGACTATAAGAGGAGCTACCCACCA 1500
QY 2768 ACTCGCTCAGAGAGCCAGCCAGCCACTCTCTCAAGAAACACACTAAGCTCCTCCA 2827
DB 1501 ACTCGCTCAGAGAGCCAGCCAGCCACTCTCTCAAGAAACACACTAAGCTCCTCCA 1560
QY 2828 ATTCTCTCACCTCCAGAAACAGAGCTTTGGCAGGGAGAGCAACCGCCCGCCGCC 2887
DB 1561 ATTCTCTCACCTCCAGAAACAGAGCTTTGGCAGGGAGAGCAACCGCCCGCCGCC 1620
QY 2888 GCGAGAGGTGGACTCATCCAGGTGACAGCTCCAGCCACTCTGSCCAGGCGGTGACTG 2947
DB 1621 GCGAGAGGTGGACTCATCCAGGTGACAGCTCCAGCCACTCTGSCCAGGCGGTGACTG 1680
QY 2948 TCTGAGGCGAGCCAGCTCAACGCTTACACTGACAAAGGTGGGGCTGAAGCGTA 3007
DB 1681 TCTGAGGCGAGCCAGCTCAACGCTTACACTGACAAAGGTGGGGCTGAAGCGTA 1740
QY 3008 GCGCTCTGCTAAAGCGGAGCTACCCCGCCCAACCAACCTTTGCTCCCTTCCACATCCA 3067
DB 1741 GCGCTCTGCTAAAGCGGAGCTACCCCGCCCAACCAACCTTTGCTCCCTTCCACATCCA 1800
QY 3068 TGAAGCCCAATGATGCTGTACATAA 3093
DB 1801 TGAAGCCCAATGATGCTGTACATAA 1826
RESULT 13
AAK94365
CQ AAK94365 standard: cDNA: 2227 BP.
XX AC AAK94365;
XX XX
D: 06-NOV-2001 (first entry)
DE Human full-length cDNA, SEQ ID NO: 3087.
XX KW Human: full length cDNA; cDNA synthesis; oligo-capping; ss.
XX OS Homo sapiens.
XX PN EPI33094-A2.
XX XX
PD 05-SEP-2001.
XX PF 07-JUL-2000; 2000EP-0114089.
XX XX
```

PR 08-JUL-1999; 95JP-0194486.  
PR 11-JAN-2000; 2000JP-0118774.  
PR 02-MAY-2000; 2000JP-0183765.  
XX  
PA (HELI-) HELIX RES INST.  
XX Ota T, Nishikawa T, Itoyai T, Hayashi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX P-PSDB: NAM93444.  
DR WP1: 2001-524255/58.  
DR  
XX 830 Primers useful for synthesizing full length cDNA clones and their  
PT use in genetic manipulation -  
XX  
XX Claim 8: SEQ ID NO 3087; 1380pp + sequence listing: English.  
XX  
XX The invention relates to primers for synthesizing full length cDNA  
CC clones. 830 cDNA molecules encoding a human protein have been  
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
CC molecules have been determined. Primers for synthesizing the full length  
CC cDNA are useful for clarifying the function of the protein encoded by  
CC the cDNA. The full length clones were obtained by construction of full  
CC length enriched cDNA libraries that were synthesised by the oligo-capping  
CC method. The primers enable the production of the full length cDNA easily  
CC without any special methods. The present sequence is a full length  
CC human cDNA of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in CD-ROM format directly from EPO.  
XX  
SQ Sequence 2227 BP: 567 A; 760 C; 566 G; 460 T; 0 other:  
  
Query Match 56.3%; Score 1742.4; DB 22; Length 2227;  
Best Local Similarity 99.9%; Pred. No. C;  
Matches 1743; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1350 TTTGGCCAGAAATAGGAATATAGTGTGTTCTAAATGACACAGCCTTTCTCGGAGGAGATGAG 1403  
Db 1 TTTGGCCAGAAATAGGNAATATAGTGTGTTCTAAATGACACAGCCTTTCTCGGAGGAGATGAG 60  
  
QY 141C TGTTTACAACTCIGAAAATCAGACTGTGATGAGTGGAGTGAAGACAAAGGATCATGGGCAT 1469  
Db 61 TGTTTACAACTCTGAAAATCAGACTGTGATGAGTGGAGTGAAGACAAAGGATCATGGGCAT 120  
  
QY 1470 GCAGTGGACAGCAGACAGCTCTCTGTATGTTGCGTTCTCTACTGTGTGATAAAGGT 1529  
Db 121 GCAGTGGACAGACAGCAGCA KCTCTCTGTATGTTGCGTTCTCTACTGTGTGATAAAGGT 130  
  
QY 1530 TCCCTTGGCCGGTGTGAAACACATGAGTGGGAAGTGTAAATAAGCTGTATTGCTCCAGAGA 1589  
Db 181 TCCCTTGGCCGGTGTGAAACACATGAGTGGGAAGTGTAAATAAGCTGTATTGCTCCAGAGA 240  
  
QY 1590 CCATATCTGGATGGATAAAGGAAGGTGGTGGCTGAGCCGATATACGACCAACACAGAG 1649  
Db 241 CCATATCTGGATGGATAAAGGAAGGTGGTGGCTGAGCCGATATATCACCACACAGAGAG 300  
  
QY 1650 ACTGACTTTTGACGAGGACATAGACGCTGGCAATACAGATGCTGTGGGGACTGTCAAA 1709  
Db 301 ACTGACTTTTGACGAGGACATAGACGCTGGCAATACAGATGCTGTGGGGACTGTCAAA 360  
  
QY 1710 TTTCCTTTGGCACTGAATGGGCATTCAGTTCCCTCTTGGCCAGCACAACCATCAGA 1769  
Db 361 TTTCCTTTGGCACTGAATGGGCATTCAGTTCCCTCTTGGCCAGCACAACCATCAGA 420  
  
QY 1770 TTGACGGCTCAAGGGGTATGAGTCTAGGCGAGGAATGTGGACTGGAAGCATCTGCT 1829  
Db 421 TTGACGGCTCAAGGGGTATGAGTCTAGGCGAGGAATGTGGACTGGAAGCATCTGCT 480  
  
QY 1830 TGACTCACTGACAGACAGACCCCTTTGGGGCAGTGTCTTCCCAATATCAACCAAGACAA 1889  
Db 481 TGACTCACTGACAGACAGACCCCTTTGGGGCAGTGTCTTCCCAATATCAACCAAGACAA 540  
  
QY 1890 GAAGGAGTGAATTCGGGAAATTAAGTCAAGGCCACGACCAAGCTGGTTCGGTCACCCCT 1949

Db 541 GAAGGAGTGAATTCGGGAAATTAAGTCAAGGCCACGACAGCTGGTTCGGTCACCCCT 600  
QY 1950 CTTGGCCATTGCAATCTGGCTTTCTGTCATGGGGCCGCTCTCTCGGGGATCACCGT 2009  
Db 601 CTTGGCCATTGCAATCTGGCTTTCTGTCATGGGGCCGCTCTCTCGGGGATCACCGT 660  
QY 2010 CTACTCGCTCTGTGATCATCGGCGCAAGACGTGTGTGTGTCAGCGCAAGGAGAAAGA 2069  
Db 661 CTACTCGCTCTGTGATCATCGGCGCAAGACGTGTGTGTGTCAGCGCAAGGAGAAAGA 720  
QY 2070 GGTCAACCACTCGCGCGGGCTCCATGAGCAGCTCACCAAGCTCAGCGGCTCTTTGG 2129  
Db 721 GGTCAACCACTCGCGCGGGCTCCATGAGCAGCTCACCAAGCTCAGCGGCTCTTTGG 780  
QY 2130 GGACACTCAATCCNAAGACCCAAAGCCGGAGGCTCATCTCAGGCCACTCATGCAACAGG 2189  
Db 781 GGACACTCAATCCNAAGACCCAAAGCCGGAGGCTCATCTCAGGCCACTCATGCAACAGG 840  
QY 2190 CAAAGTCCGACCTCCCGGCAACACGCGCAAGATGCTCATTAAGCAGACGACCACTCT 2249  
Db 841 CAAAGTCCGACCTCCCGGCAACACGCGCAAGATGCTCATTAAGCAGACGACCACTCT 900  
QY 2250 GGACCTGAGCGGCTCCCGGCAACACGCGCAAGATGCTCATTAAGCAGACGACCACTCT 2309  
Db 901 GGACCTGAGCGGCTCCCGGCAACACGCGCAAGATGCTCATTAAGCAGACGACCACTCT 960  
QY 2310 CAGCCGCGGAGCGCGGAGTGGAGAGAACAGACACTCATCAATGCTGCACAAAGA 2369  
Db 961 CAGCCGCGGAGCGCGGAGTGGAGAGAACAGACACTCATCAATGCTGCACAAAGA 1020  
QY 2370 CATGCCCGCCATGGGCTCCCTGTGATTCGCCAGGACCTGCCCTGGCGGCTCCCGCAG 2429  
Db 1021 CATGCCCGCCATGGGCTCCCTGTGATTCGCCAGGACCTGCCCTGGCGGCTCCCGCAG 1080  
QY 2430 CCACATCCCGCAGGCTGTCTCTGCCATCAGCAGCAGGCTACCAACATGAGTACCT 2489  
Db 1081 CCACATCCCGCAGGCTGTCTCTGCCATCAGCAGCAGGCTACCAACATGAGTACCT 1140  
QY 2490 GGACGAGCCCAATGAGCGAGTGGCCGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2549  
Db 1141 GGACGAGCCCAATGAGCGAGTGGCCGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
QY 2550 GGAGTATAAGACCATCAAGGAACTCTCAGCAGCAAGTGTCCCAACCATGGGGTGAACCT 2609  
Db 1201 GGAGTATAAGACCATCAAGGAACTCTCAGCAGCAAGTGTCCCAACCATGGGGTGAACCT 1260  
QY 2610 TGTGGAGAGCTGGACAGCTGCCCGGCAAGTTCACAGCGGAGGAGGAGGAGGAGGAGGAG 2669  
Db 1261 TGTGGAGAGCTGGACAGCTGCCCGGCAAGTTCACAGCGGAGGAGGAGGAGGAGGAGGAG 1320  
QY 2670 CCGGGAGGCTTCCTGTCTCAGAGCTGGTCTAAGCAAGCGGCTGGAATGCACGACTCTCT 2729  
Db 1321 CCGGGAGGCTTCCTGTCTCAGAGCTGGTCTAAGCAAGCGGCTGGAATGCACGACTCTCT 1380  
QY 2730 TTCTTAGGGGTGTGACTATTAAGAGGAGCTACCCGCAAGTGTCTCAGGAGGAGGAGGAGGAG 2789  
Db 1381 TTCTTAGGGGTGTGACTATTAAGAGGAGCTACCCGCAAGTGTCTCAGGAGGAGGAGGAGGAG 1440  
QY 2790 GGCACACACTCTCAAGAGAAACAACTAACTCTCTCAATTCCTCTCAGCTCTCCAGAAA 2849  
Db 1441 GGCACACACTCTCAAGAGAAACAACTAACTCTCTCAATTCCTCTCAGCTCTCCAGAAA 1500  
QY 2850 CCAGAGCTTTTGGCAGGGGAGACAAACCGCGCCGCGCCGCGCAGAGGGTGGACTTCCATCA 2909  
Db 1501 CCAGAGCTTTTGGCAGGGGAGACAAACCGCGCCGCGCCGCGCAGAGGGTGGACTTCCATCA 1560  
QY 2910 GGTGCACAGCTCCGACCATCTCGCCAGGCGGTGACTGTCTCGAGGAGCAGCCGACCTCAA 2969  
Db 1561 GGTGCACAGCTCCGACCATCTCGCCAGGCGGTGACTGTCTCGAGGAGCAGCCGACCTCAA 1620  
QY 2970 GCGCTACAACTCACTGACRAGGTGCGGGCTGAAGCTAGCGCTCGCTCAAGCGGAGCT 3029



Db	1126	TCCCTCTGTGATTCCACGGAGACTGCCCTTGC	CGGUCCTCCGCCAGCCACATCCACGCGTG	1188
QY	2446	GTGCTCTGTGCCATCATCAGCAGCAGCGCTAC	CAAGCACTACCAAGCACTAGATAGCTGGACAC	2505
Db	1186	GTGCTCTGTGCCATCATCAGCAGCAGCGCTAC	CAAGCACTACCAAGCACTAGATAGCTGGACAC	1245
QY	2506	ACGAGGTGGCCCATATGGCGCTGGAGAGACAG	CGCGCCACACTCTGGAGTATAGACCATC	2565
Db	1246	ACGAGGTGGCCCATATGGCGCTGGAGAGACAG	CGCGCGCCACACTCTGGAGTATAGACCATC	1305
QY	2566	AAGGAACATCTCAGCAGCAAGTCCCAACCACT	GGGGTCAACCTTGTTGAGAACTCTGCAC	2625
Db	1306	AAGGAACATCTCAGCAGCAAGTCCCAACCACT	GGGGTCAACCTTGTTGAGAACTCTGCAC	1365
QY	2626	AGCTTGGCCCCCAAGTTCTCACAGCGGGAGGC	CTCCCTTGGGTTCGCCCGGAGCGCTTCCTG	2685
Db	1366	AGCTTGGCCCCCAAGTTCTCACAGCGGGAGGC	CTCCCTTGGGTTCGCCCGGAGCGCTTCCTG	1425
QY	2686	TCTCAGACCGGCTTCAAGCAACGCGCTCGAA	TCCACCTCTCTCTACGGGCTTGAC	2745
Db	1426	TCTCAGACCGGCTTCAAGCAACGCGCTCGAA	TCCACCTCTCTCTACGGGCTTGAC	1485
QY	2746	TATAAGAGAGCTATGCCACATTAACCTGCTCA	CAAGCAAGGCAACAGGCACTCTCAAA	2805
Db	1486	TATAAGAGAGCTATGCCACATTAACCTGCTCA	CAAGCAAGGCAACAGGCACTCTCAAA	1545
QY	2806	AGAACACACTAATCTCTCTCTCTCTCTCTCT	CTCAGAACAGAGGCTTTGSCAG	2865
Db	1546	AGAACACACTAATCTCTCTCTCTCTCTCTCT	CTCAGAACAGAGGCTTTGSCAG	1605
QY	2866	GGAGACAACCGCGCGCGCGCGCGCGCGCGCG	CGAGAGGTTGGACTTCATCAGAGTCACAGTCC	2925
Db	1606	GGAGACAACCGCGCGCGCGCGCGCGCGCGCG	CGAGAGGTTGGACTTCATCAGAGTCACAGTCC	1665
QY	2926	CCATCTGGCAGCGCTGACGTCTGAGGCGAGCC	AGCCCTCAAGCGCTTACAGTCACTG	2985
Db	1666	CCATCTGGCAGCGCTGACGTCTGAGGCGAGCC	AGCCCTCAAGCGCTTACAGTCACTG	1725
QY	2986	ACAAGTCTGGGCTTGAAGCGTACGCGCTGGCT	TAAGCGGAGGTACCCCGCAAAACCATCC	3045
Db	1726	ACAAGTCTGGGCTTGAAGCGTACGCGCTGGCT	TAAGCGGAGGTACCCCGCAAAACCATCC	1785
QY	3046	TTTGTCGCCCTTCCATCATCATCATCATCATCAT	GTGGTGTCATATA	3093
Db	1786	TTTGTCGCCCTTCCATCATCATCATCATCATCAT	GTGGTGTCATATA	1833
RESULT 15				
AAS68253				
ID	AAS68253 standard; cDNA; 2592 BP.			
AC	AAS68253;			
CC				
DT	13-FEB-2002 (first entry)			
DE	DNA encoding novel human diagnostic protein #4057.			
XX				
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;			
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.			
XX	Homo sapiens.			
XX				
PN	W0200175067-A2.			
XX				
PD	11-OCT-2001.			
XX				
PF	30-MAR-2001; 2001WO-US08631.			
XX				
PR	31-MAR-2000; 2000US-0540217.			
XX				
PR	23-AUG-2000; 2000US-0649167.			
XX				
PA	(HYPE-) HYSEQ INC.			
XX				

Drmanac KI, Liu C, Tang YT.  
 WPI: 2001-639362/73.  
 P-PSDB; ABG04066.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -

Claim 1: SEQ ID NO 4057; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

Sequence 2592 BP; 655 A; 754 C; 659 G; 524 T; 0 other;

Query Match 51.5%; Score 1592; DB 25; Length 2592;  
 Best Local Similarity 84.3%; Pred. No. 0;  
 Matches 1809; Conservative 0; Mismatches 325; Indels 12; Gaps 1;

960 CAGCATCCCGGGGTGCGAGTGTGGCTATGACATGCTTGACATGCCAGTGTGTTTAC 1019  
 |||||  
 447 CAGATCCCTGGTGTCTCGATCTGTGGCTATGACATGCTTGACATGCCAGTGTGTTTAC 506  
 |||||

1020 TGGGAGATTCAAGGACACAGAAAGTCTCTGATTCACCTGGACACCACTGCTGAIGAAG 1079  
 |||||  
 507 TGGGAGATTCAAGGACACAGAAAGTCTCTGATTCACCTGGACACCACTGCTGAIGAAG 566  
 |||||

1080 AGTTCCTTAAGCCAGGCCAGGTCCTCTGTGGTCTCATCTCTCTTACAAGATATGCAAC 1139  
 |||||  
 567 AGTTCCTTAAGCCAGGCCAGGTCCTCTGTGGTCTCATCTCTCTTACAAGATATGCAAC 626  
 |||||

1140 CTCCAATGAGTTCCTCATATGATACCTTGAACTTCATCAAGACGCCACCGCTCATGGATGA 1199  
 |||||  
 627 CTCCAATGAGTTCCTCATATGATACCTTGAACTTCATCAAGACGCCACCGCTCATGGATGA 686  
 |||||

1200 GGCAGTGCCTTCATCTCTCAACAGGCCAATGGTTCCTGAGACAAATGTCAGATACCGCT 1259  
 |||||  
 687 GGCAGTGCCTTCATCTCTCAACAGGCCAATGGTTCCTGAGACAAATGTCAGATACCGCT 746  
 |||||

1260 TACCAAAATTGCAGTGGACACAGCTGCTGGGCCATATCAGAATCACACTGTGGTGTTCCT 1319  
 |||||  
 747 TACCAAAATTGCAGTGGACACAGCTGCTGGGCCATATCAGAATCACACTGTGGTGTTCCT 806  
 |||||

1320 GGCATCAGAAAGGAAATCATCTTTGAAGTATTTTGGCCAGAAATAGGAATAAGTGGTTCCT 1379  
 |||||  
 807 GGCATCAGAAAGGAAATCATCTTTGAAGTATTTTGGCCAGAAATAGGAATAAGTGGTTCCT 866  
 |||||

1380 AATGACAGCCTTTTCTCGAGGAGATGAGTGTTTACAACCTCTGAAAAATGCAGCTATGA 1439  
 |||||  
 867 AATGACAGCCTTTTCTCGAGGAGATGAGTGTTTACAACCTCTGAAAAATGCAGCTATGA 926  
 |||||

1440 TGGAGTCCGAGACAAAAGGATCATGCGCATGCGGCATGCGACAGACCAAGCAGCTCTCTGTA 1499  
 |||||

[illegible]

Db	2007	CGAGGTGGCCAGATGGCGTGTGAGAGACCAAGCGCCGACACTGGAGCTATAAGACCAATCAA	2066
Qy	2568	GGAACTATCTCAGCAGCANGAGTGTCCCAACCATGGGGTGAACCTTGTGAGAAAGCTGGACAG	2627
Db	2067	GGAACTATCTCAGCAGCANGAGTGTCCCAACCATGGGGTGAACCTTGTGAGAAAGCTGGACAG	2126
Qy	2628	CGTGCCTCCGCAAAAGTTCACAGCGGGAGGCGCTCCCTGSGGTGCCCGCGAGACCTCCCTGTG	2687
Db	2127	CGTGCCTCCGCAAAAGTTCACAGCGGGAGGCGCTCCCTGSGGTGCCCGCGAGACCTCCCTGTG	2186
Qy	2688	TCAGACCGGTCTTAAGCAAGCGCTGGAATGCAACCACTCCTCTCTCTACGGGGTTCACCTA	2747
Db	2187	TCAGACCGGTCTTAAGCAAGCGCTGGAATGCAACCACTCCTCTCTCTACGGGGTTCACCTA	2246
Qy	2748	TAAAGAGAGCTACCCCAAGCAAGTCTGCTCAGGAGAGCCACAGCCAGCACTCTCAAAAG	2807
Db	2247	TAAAGAGAGCTACCCCAAGCAAGTCTGCTCAGGAGAGCCACAGCCAGCACTCTCAAAAG	2306
Qy	2808	AAACAACACTAATCTCTTCCAAATTCCTCTCACTCTCTCCAGAAACAGAGTTTGGCAGGGG	2867
Db	2307	AAACAACACTAATCTCTTCCAAATTCCTCTCACTCTCTCCAGAAACAGAGTTTGGCAGGGG	2366
Qy	2868	AGACAACCGCCGCGCCGCGCAGAGAGGCTGACTCCATCAGGTGCACAGCTCCACAGC	2927
Db	2367	AGACAACCGCCGCGCCGCGCAGAGAGGCTGACTCCATCAGGTGCACAGCTCCACAGC	2426
Qy	2928	ATCTGCGCAGGCGGTGACTGTCTCTCGAGGACGCCAGCGCTCAAGCGCTACAACTCACTGCAC	2987
Db	2427	ATCTGCGCAGGCGGTGACTGTCTCTCGAGGACGCCAGCGCTCAAGCGCTACAACTCACTGCAC	2486
Qy	2988	AAGGTGGGGCTGAGCGTACGCCCTCGCTTAAGCGCGAGCTACCCCGCAAAACCATCCCT	3047
Db	2487	AAGGTGGGGCTGAGCGTACGCCCTCGCTTAAGCGCGAGCTACCCCGCAAAACCATCCCT	2546
Qy	3048	TGCTCCCGCTTCCACATCCCATGAAGCCCAATGATGGGTGTACATAA	3093
Db	2547	TGCTCCCGCTTCCACATCCCATGAAGCCCAATGATGGGTGTACATAA	2592

Search completed: September 30, 2003, 16:58:34  
Job time : 808.168 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2003, 16:35:33 : Search time 5937.36 Seconds  
(without alignments)  
12661.132 Million cell updates/sec

Title: US-09-856-681-1

Perfect score: 3093

Sequence: 1 atgaggtcaagagcttgct.....ccatgctgctgtacataa 3093

Scoring table:

IDENTITY\_NJC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estm.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_hic.\*

9: gb\_estl.\*

10: gb\_est2.\*

11: gb\_est3.\*

12: gb\_est4.\*

13: gb\_est5.\*

14: gb\_est6.\*

15: em\_estin.\*

16: em\_estom.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_piv.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_mam.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_rod.\*

26: em\_gss\_phg.\*

27: em\_gss\_vrl.\*

28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3080.8	99.9	6875	11	BC032619 Homo sapi
2	1439.4	46.5	3226	11	AK042751 Mus muscu
3	1437.6	46.5	3329	11	AK082711 Mus muscu
4	1065.6	34.5	2411	14	CB605722 AMGNNUC:M

5	899.4	29.1	1030	12	BM450002
6	896	29.0	1183	12	BM546059
7	828.2	26.8	868	9	AU140366
8	813.8	26.3	891	13	B0186963
9	784.4	25.4	887	12	BG769297
10	783.4	25.3	1201	9	AL543344
11	779.4	25.2	848	13	BQ578336
12	772.4	25.0	883	13	B0172225
13	770	24.9	864	13	B0440312
14	767.4	24.8	874	13	B0838082
15	760.5	24.6	890	13	B0683009
16	757.4	24.5	953	13	B0854884
17	734	23.7	952	13	B0855855
18	713.8	23.1	835	10	BG477592
19	682.2	22.1	4374	11	AK031307
20	661.8	22.0	3921	11	AK084922
21	680.2	22.0	4476	11	AK052232
22	670.2	21.7	716	10	BF970807
23	667	21.6	890	13	B0856543
24	664.8	21.5	736	10	HE277845
25	660.4	21.4	794	10	AW954605
26	656.2	21.2	823	10	BG327694
27	645.8	20.9	651	9	AL502452
28	636.6	20.6	643	10	HE408781
29	631.8	20.4	685	10	HE265000
30	629.8	20.4	703	10	HE384511
31	615.2	19.9	926	10	BG326467
32	614.4	19.9	746	13	BQ770491
33	614.4	19.9	795	14	CB245479
34	612.2	19.8	772	13	B0425051
35	611.8	19.8	895	12	B1819955
36	608.8	19.7	777	13	B0708565
37	604.4	19.5	771	14	CA513024
38	589.6	19.1	829	12	B1824613
39	589.2	19.0	690	10	BG33712
40	588.8	19.0	592	14	CA337090
41	588.6	19.0	884	10	HE728944
42	578	18.7	603	10	BG334875
43	576.8	18.6	852	10	BF306161
44	573	18.5	713	10	BF389749
45	570.4	18.4	612	10	BG327283

## ALIGNMENTS

RESULT 1  
BC032619  
LOCUS BC032619  
DEFINITION Homo sapiens, similar to sema domain, transmembrane domain (TM) and cytoplasmic domain. (sexaphorin) 6A, clone IMAGE:5578065, mRNA.  
ACCESSION BC032619  
VERSION BC032619.1  
KEYWORDS HT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 6875)  
AUTHORS Strausberg,R.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov  
COMMENT Contact: MGC help desk  
Email: cgabs@mail.nih.gov  
Tissue Procurement: ATCC/DCID/BTP  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),

BC032619 6875 bp mRNA linear HTC 04-MAR-2003  
Homo sapiens, similar to sema domain, transmembrane domain (TM) and cytoplasmic domain. (sexaphorin) 6A, clone IMAGE:5578065, mRNA.

BC032619.1 GI:22749800  
HTC.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 6875)

Direct Submission

Submitted (06-JUN-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: http://mgc.ncl.nih.gov

Contact: MGC help desk

Email: cgabs@mail.nih.gov

Tissue Procurement: ATCC/DCID/BTP

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland:  
 Web Site: <http://www.nisic.nih.gov/>  
 Contact: [nisc.mcgm@nih.gov](mailto:nisc.mcgm@nih.gov)  
 Akhter, N., Ayello, K., Beckstrom-Stanberg, S.M., Benjamin, B.,  
 Binkley, K.W., Bouffard, G.O., Brown, K., Brinkley, C.C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghghi, P.,  
 Hanson, N., Ho, L., Karlis, E., Kwong, P., Lari, P., Legasi, R.,  
 Maduro, Q. L., Mastello, C., Masker, B., Mastrian, S.J., McCloskey, J. C.,  
 McDowell, J., Pearson, R., Skenteripop, S., Thomas, P.J., Touchman, W.,  
 Tusarguen, C., Voigt, J.L., Walker, M.A., Welch, K.S., Wilgus, L.,  
 Young, A., Zhang, L.H. and Green, E.P.

Clone distribution: M3C clone distribution information can be found through the I.M.A.G.E. Consortium/LLN at: <http://image.lln.gov>  
Series: IRAK Plate: 69 Row: 6 Column: 4  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA at: 11991459  
This clone has the following problem: retained intron.

[illegible]

```

1: 5875
/organism="Homo sapiens"
/mol_type="mRNA"
/CD_xref="taxon:9606"
/clone_image="5578366"
/clone_type="Sk10, melanot
/clone_lib="NH_MGC_72"
/lab_host="pHL03"
/note="Vector: pCMV-Sp666"
1784 a 1795 c 182 g 1704 t

```

BASE COUNT	1784 a	1795 c	1592 c	1704 t
ORIGIN				

Query Match 99.9%; Score 3089.8; DB 11; Length 6875;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3091; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATCAGGTCAGAACCCCTTGCTGCTATATATTCACACTGCTACACTTTCGCTGGGCTGGTTC	60
Db	792	ATGAGGTCAGAACCCCTTGCTGCTATATTCACACTGCTACACTTTCGCTGGGCTGGTTC	851
QY	61	CCAGAAGATCTTCAGGCCAATCAGATATTCGCA*GGCAATATACAAACAGTATCGGCTG	120
Db	852	CCAGAAGATCTTCAGGCCAATCAGATATTCGCAATGCGCACTATACAAACAGTATCGGCTG	911
QY	121	TTTGTGGGTCACAAGCCAGGACGGSACACACACACAGCCACAGGCTGGACATCCAGATG	180
Db	912	TTTGTGGGTCACAAGCCAGGACGGSACACACACACAGG*ACAGGCTGGACATCCAGATG	971
QY	181	ATTATGATCATGAACGGAGACCCCTACATTCGCTGCTAGGACCATAT*TTATATGTTGAT	240
Db	972	ATTATGATCATGAACGGAGACCCCTACATTCGCTGCTAGGAGCAATATTTATATGTTGAT	1031
QY	241	ATAGACACATCAGACACAGCGGAAGAAATTTATG*AGCAAAAACCTGCATGSAATCTAGA	300
Db	1032	ATAGACACATCAGACACAGCGGAAGAAATTTATG*AGCAAAAACCTGCATGSAATCTAGA	1091
QY	301	CAGCGCGATCTACACACATCCAGAAATCGAGGGAACAT*TAAGCATGATAGTCCACACTTT	360
Db	1092	CAGCGCGATCTAGACATGCGAATCGAGAAATGAAGGGAACATAGGATGATGCCACACTTT	1151
QY	361	ATTAAAGTCTTCTAAGAAACAGATGATGCATTTTGTGCTGSAATTAATGCTTC	420
Db	1152	ATTAAAGTCTTCTAAGAAACAGATGATGCATTTTGTGCTGSAACATAATGCTTC	1211
QY	421	AACCCCTCCCTCGAGAACTATAGATGGATACATTCGAACCA*TCGCGATGAATTCAGC	480
Db	1212	AACCCCTCCCTCGAGAACTATAGATGGATACATTCGAACCA*TCGCGATGAATTCAGC	1271
QY	481	GGAA*GGCCAGATGCCCATATGATGCCCAACATGCCAACCTTCGCACTGTTTCAGATGGA	540
Db	1272	GGAA*GGCCAGATGCCCATATGATGCCCAACATGCCAACCTTCGCACTGTTTCAGATGGA	1331
QY	541	AAACTATATCTACGCCACAGTACTGACTTCCTTGCCATTCGACCATTCATTTACCGAGAT	600



27-51	CCACGAACTCGGTCAAGAGAACCACAGGCCACCACTTCTCAAAGAAACAACACTAAC	2892
DB		
3552	CCACGAACTCGGTCAAGAGAACCACAGGCCACCACTTCTCAAAGAAACAACACTAAC	3611
QY		
2821	TCTTGGAATTCCTCTACCTCTCCAGNAACCAGAGGTTGGCAGGGGAGACAACCCGC	2883
DB		
3612	TCTTCCAAITCCTCTCACCTCTCCAGAAACCAGAGCTTGGCAGGGGAGACAACCCGC	3671
QY		
2881	CCGCCCCGAGAGSGTGAATCCATCCAGGTGCACAGCTCCCAGCATCTTGGCCAGGCC	2940
DB		
3672	CCGCCCCGAGAGSGTGAATCCATCCAGGTGCACAGCTCCCAGCATCTTGGCCAGGCC	3731
QY		
2941	GIGACTGTCTCGAGCGAGCCAGCCAGCTCAAGCGCTACAACTCAGTCAACAGGTCTCGGGCTG	3000
DB		
3732	GIGACTGTCTCGAGCGAGCCAGCCAGCTCAAGCGCTACAACTCAGTCAACAGGTCTCGGGCTG	3791
QY		
3001	AAGGTGATCGCCTCGCTTAAGAGCGGAGCTACCCGCCAAAACCATCTTTGCTCCCTTTCC	3060
DB		
3792	TAGGTAGCGCCCTCGCTTAAGAGCGGAGCTACCCGCCAAAACCATCTTTGCTCCCTTTCC	3851
QY		
3061	ACATCCATGAAGCCCAATGATGGGTGATACATAA	3093
DB		
3852	ACATCCATGAAGCCCAATGATGGGTGATACATAA	3884

  

RESULT 2	AK042751	AK042751	3226 bp	mRNA	linear	HTC 05-DEC-2002
LOCUS	DEFINITION	Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length				
VERSION		enriched library, clone:A730020P05 product:sema domain,				
KEYWORDS		transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A,				
SOURCE		full insert sequence.				
ORGANISM	ACCESSION	AK042751				
	VERSION	AK042751.1	GI:26335300			
	KEYWORDS	HTC; CAP trapper.				
	SOURCE	Mus musculus (house mouse)				
	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
	AUTHORS	Carninci,P. and Hayashizaki,Y.				
	TITLE	High-efficiency full-length cDNA cloning				
	JOURNAL	Meth. Enzymol. 403, 19-44 (1999)				
	REFLINE	99279253				
	PUBMED	10349636				
	REFERENCE	2				
	AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,				
	TITLE	Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.				
	JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to				
	REFLINE	prepare full-length cDNA libraries for rapid discovery of new genes				
	PUBMED	Genome Res. 10 (10), 1617-1630 (2000)				
	REFERENCE	20495374				
	AUTHORS	-1042159				
	TITLE	HIKEN integrated sequence analysis (RISA) system--384-format				
	JOURNAL	sequencing pipeline with 384 multicapillary sequencer				
	REFLINE	Genome Res. 10 (11), 1757-1771 (2000)				
	PUBMED	20530913				
	REFERENCE	11076861				
	AUTHORS	Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,				
	TITLE	Arakawa,T., Hara,A., Fukunishi,Y., Kono,H., Adachi,J., Fukuda,S.,				
	JOURNAL	Saito,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamahaka,I.,				
	REFLINE	Salto,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,				
	PUBMED	Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,I.,				
	REFERENCE	Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kichiwa,H.,				
	AUTHORS	Shibata,K., Itoh,M., Alizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P.,				
	TITLE	Kono,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,				
	JOURNAL	Sum,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,				
	REFLINE	Yamamoto,R., Matsumoto,H., Sakauchi,S., Ikegami,I., Kashiwagi,K.,				
	PUBMED	Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wachiaki,M.,				
	REFERENCE	Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,				
	AUTHORS	Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.				

Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pessiole, G.,  
Quackenbush, J., Schriml, M., Staebli, F., Suzuki, K., Tomita, M.,  
Wagner, I., Washio, I., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., De Bonaldo, M. F., Brownstein, M. J., Bull, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustiguchi, S., Hill, D.,  
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Marzari, J., Mombaerts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, J., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seyar, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, G., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S.,  
and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821): 695-698 (2001)

21045660

11217851

5

The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II team

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420: 563-573 (2002)

6 (bases 1 to 3226)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, P., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiracka, T., Hirozane, T.,  
Hori, F., Imotani, K., Isuli, Y., Iton, M., Kageoka, I., Kasakawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, K.,  
Koya, S., Kurikawa, C., Matsuyama, I., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Nomazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, K., Saito, K., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki-Akahira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, I., Yasunishi, A.,  
Yamamatsu, M., and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN,  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://phantom.gsc.riken.go.jp/

URL: http://phantom.gsc.riken.go.jp/

FEATURES

source

1..3226  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_05:A73020P05"  
/db\_xref="taxon:10090"  
/clone="A73020P05"  
/tissue\_type="cerebellum"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="7 days neonate"  
671..2488

CDS

/note="unnamed protein product; putative  
scd4 domain, transmembrane domain (TM), and cytoplasmic  
domain, (senaphorin) 5A (MG0\_MG7:1203727, GRINM\_015744,  
evidence: BLASTN, 100%, match=3402)."  
/codon\_start=1  
/protein\_id="BAC31351.1"  
/db\_xref="EG:2433530"  
/translation="KRPAALLLCLTLICACAGAFPHDSPPISISHNVTQYVPVGH  
KFGNRTQHRLLDLCIMIMNNTLVVAARDHIYIVDIDTSHTEEYQCSKSLTKWSRQA

DVVICPMKCKHKHDECHNHFKVLLKKNDQTLFVCGTNAENPSCRNVRVDITLFTGDEFS  
CHACPYDAKHANIALPADCKLYSATVDFDLAIDAVYRSLCDSTPLRVHDSKWLK  
EPFVQAVDGTSTIFFREIAETVIRKGVVPRVQVCKNDMSGSRVLEKQWTS  
LAKRLNCSVPGSHFYFNLTAVTVIRNGRVVLTATSTPYNIPISVACAYDMID  
IANVTGREAFKQSPDSTWTPVDFPKPRPGCCAGSSGLEKYATSNFPDDTLNFI  
KTIPLIMDEAVPSTINRPMFLRTKVAVYRLTKIAVDNAGPYQNHVTVPFLGSEKGLIKF  
LARISSGFLNGLSLEEMNVTNPKCSYDGVGDKRIMGDMOLDRASGLSYVAFSTCVI  
KVLPSCEHCKKCTCIASRDYPCWVRESGCAHLSPLSRLTFEODIERNGTDGLG  
KCHNSFVALNDISTPLPDYKSEFNRYGFIATYIYIN"

BASE COUNT 883 a 752 c 731 g 860 t

CPUSIN

Query Match 46.5%; Score 1439.4; DB 11; Length 3226;

Best Local Similarity 89.1%; Pred. No. 0;

Matches 1554; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 1 ATGAGGTCAAGAGCCITGCTGTATATTTACACTGCTACATTTGCTGGGCTGGTTTC 60

DE 671 ATGGGCCACAGCCTTACTGCTGTCTCACACTGCTACACTGGCGCTGGTTTC 730

QY 61 CCAGAGATCTCTGAGCAATCAGTATTTCCGATGGCACTATACAAACAGTATCGGTG 120

DB 731 CCAGAGATCTCCGAGCAATCAGTATTTCCGATGGCACTATACAAACAGTATCGGTG 790

QY 121 TTITGGGCGCAAGCCAGGACACACACAGAGAGAGGCTGACATCCAGATG 180

DB 791 TTITGGGCGCAAGCCAGGACACACACAGAGAGAGGCTGACATCCAGATG 850

QY 181 ATTATGATCATGACGGAGCCCTCTACATGCTGTAGGACATATTTATCTGTTGAT 240

DB 851 ATCATGATCATGACAGAGAGCCCTCTACATGCTGTAGGACATATTTATCTGTTGAT 910

QY 241 ATAGACATATCAGACAGGAGAAATTTATTGTAGCAAAAACATGACATGGAATCTAGA 300

DB 911 ATAGACATATCAGACAGGAGAAATTTATTGTAGCAAAAACATGACATGGAATCTAGA 970

QY 301 CAGGCGATGTATACATGATGAGATGAGAGGAGAAACATAGAGTGTGCGCACAACTTT 360

DB 971 CAGGCGATGTATACATGATGAGATGAGAGGAGAAACATAGAGTGTGCGCACAACTTT 1030

QY 361 ATTAAGTTCTTCTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 420

DB 1031 ATTAAGTTCTTCTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1090

QY 421 AACCTTCTCTGAGAACTATAAGATGATGATGATGATGATGATGATGATGATGATGATG 480

DB 1091 AACCTTCTCTGAGAACTATAAGATGATGATGATGATGATGATGATGATGATGATGATG 1150

QY 481 GGAATGCCAGATGCCCATATGATGATGATGATGATGATGATGATGATGATGATGATG 540

DB 1151 GGAATGCCAGATGCCCATATGATGATGATGATGATGATGATGATGATGATGATGATG 1210

QY 541 AAATCTACTCAGCCACAGTACTGATGATGATGATGATGATGATGATGATGATGATGATG 600

DB 1211 AAATCTACTCAGCCACAGTACTGATGATGATGATGATGATGATGATGATGATGATGATG 1270

QY 501 CTGGAGAGAGCCCTACCTGCGGAGCGCTGACAGCAGATCAAAATGGTTTGAAGAACCA 560

DB 1271 CTGGAGAGAGCCCTACCTGCGGAGCGCTGACAGCAGATCAAAATGGTTTGAAGAACCA 1330

QY 561 TACTTTCTCAAGCCCTGGATACGGAGATATATCTACTTCTTCAGGGAATAGCA 720

DB 1331 TACTTTCTCAAGCCCTGGATATACGGAGATATATCTACTTCTTCAGGGAATAGCA 1390

QY 721 GTGAGATTAACACCATGGGAAGGTAGTTTCCCAAGAGTGGCTCAGGTTTGTAGAAT 780

DB 1391 GTGAGATTAACACCATGGGAAGGTAGTTTCCCAAGAGTGGCTCAGGTTTGTAGAAT 1450

QY 781 GATATGGAGGATCTCAAAGAGTCCCTGGGAGAACAGTGGACGCTGTTCTCGAAGCGCC 840

DB 1451 GATATGGAGGATCTCAAAGAGTCCCTGGGAGAACAGTGGACGCTGTTCTCGAAGCGCC 1510

QY 841 TTGAATGCTCTAGTCTCTGAGAGACTCTCATTTTTTATTTTCAACATTCCTCCAGGCA 900

```

Db      1511 CTGACGTCGCTGCTGAGACTGCATTATTTCATATTAATCAATCACTCCAGCGACATACA 1570
Qy      901 GATGTGATTCGTATCAAGCGGCGTATGTTGTCCTGGCAACGTTTCTACACGCTATATAAC 960
Db      1571 GATCTGATTCCGATTATATGCGCGTATGTTGTCCTGGCAACGTTTCTACACGCTATATAAC 1630
Qy      961 AGCATCCCTGGGCTCGACGCTGTGCGCTATGACATGCTTGGACATGCGCAGTGTCTTTACT 1020
Db      1631 AGCATCCCGAGTCTCGACGCTGTGCGCTATGACATGCTTGGACATGCGCAGTGTCTTTACT 1690
Qy      1021 GCGAGATTCAGAGCAAGAGCTCTCTGATTCACATGCGACACACCTCTCTCTGATGAAGCA 1080
Db      1691 GCGAGGTTCAAGCAAGCAAGATCACTGACTCTCTCTGACACACCTCTCTCTGATGAAGCA 1750
Qy      1081 GTTCTTAAGCCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db      1751 GTCTCTAAGCCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1810
Qy      1141 TCCAAATGAGTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
Db      1811 TCCAAATGAGTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1870
Qy      1201 GCAGTGGCTTCCATCTTCACAGGCGCATGCTGCTGAGAACATGCTGATGATGATGATGATGATG 1260
Db      1871 GCAGTGGCTTCCATCTTCACAGGCGCATGCTGCTGAGAACATGCTGATGATGATGATGATGATG 1930
Qy      1261 ACCAAATGAGTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db      1931 ACCAAATGAGTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1990
Qy      1321 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
Db      1991 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2050
Qy      1381 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
Db      2051 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2110
Qy      1441 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db      2111 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2170
Qy      1501 GTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Db      2171 GTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2230
Qy      1561 TGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Db      2231 TGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2290
Qy      1621 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Db      2291 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2350
Qy      1681 AATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db      2351 AATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2410
Qy      1741 TCCCT 1745
Db      2411 CCTCT 2415

RESULT 3
LOCUS AK082711
DEFINITION Mus musculus 3 day neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:C23009A19 product:sema domain,
transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A,
full insert sequence.
ACCESSION AK082711
VERSION AK082711.1 GI:26349884;

```

```

KEYWORDS HTC: CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 Carninci,P. and Hayashizaki,Y.
          High-efficiency full-length cDNA cloning
          Meth. Enzymol. 303, 19-44 (1999)
          93279253
REFERENCE 2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
          Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
          Normalization and subtraction of cap-trapper-selected cDNAs to
          prepare full-length cDNA libraries for rapid discovery of new genes
          Genome Res. 10 (10), 1617-1630 (2000)
          20499374
REFERENCE 3 Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P.,
          Konno,H., Akiyama,Y., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
          Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
          Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
          Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,K.,
          Oyeda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
          Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
          RIKEN integrated sequence analysis (RISA) system--384-format
          sequencing pipeline with 384 multicapillary sequencer
          Genome Res. 10 (11), 1757-1771 (2000)
          20530913
REFERENCE 4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
          Arakawa,K., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
          Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
          Saito,T., Okazaki,Y., Gojohori,T., Bono,H., Kasukawa,T., Saito,R.,
          Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
          Fleischmann,K., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
          Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Suzuki,R.,
          Quackenbush,J., Schiraldi,L.M., Staubli,F., Suzuki,R., Tomita,M.,
          Warner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
          Baldracci,L.R., Parsl,G., Blake,J., Boffelli,P., Bojunga,N.,
          Carninci,P., de Boraldo,M.F., Brownstein,M.J., Bult,C.,
          Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
          Hofmann,M., Hume,D.A., Kaniya,M., Lee,N.H., Lyons,P.,
          Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
          King,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
          Sato,K., Schombach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
          Toyooka,K., Wang,K.H., Weitz,C., Whitaker,C., Wilming,L.,
          Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohsaki,S.,
          and Hayashizaki,Y.
          Functional annotation of a full-length mouse cDNA collection
          Nature 409 (6821), 685-690 (2001)
          21085660
REFERENCE 5 The PANTOM Consortium and the RIKEN Genome Exploration Research
          Group Phase I & II Team.
          Analysis of the mouse transcriptome based on functional annotation
          of 60,770 full-length cDNAs
          Nature 420, 563-573 (2002)
          1217851
REFERENCE 6 (bases 1 to 3329)
          Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
          Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
          Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
          Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
          Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
          Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
          Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
          Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
          Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
          Sogabe,Y., Tagami,M., Tadawa,A., Takahashi,F., Takaku-Akaira,S.,

```

Takeda, Y., Tsuchida, T., Tomaru, A., Taya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission:  
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Exploration: Research Group, RIKEN Genomic Sciences Center, RIKEN Yokohama Institute, 1-7-22 Suchoiro-cho, Tsurumi-ku, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp); URL: <http://genome-disc.riken.go.jp/>; Tel.: +81-45-503-9222; Fax: +81-45-503-9216

COMMENT

cdna library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our website for further details.

URL: <http://genome.qsc.riken.go.jp/>  
URL: <http://fantom.qsc.riken.go.jp/>

FEATURES	SOURCE
----------	--------

```

location/Qualifiers
1. .3329
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:C230694A19"
/db_xref="taxon:10090"
/clone="C23034A19"
/tissue_type="cerebellum"
/clone_lib="Riken full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
673..2436

```

SDS

/note="unnamed protein product; putative sema domain, transmembrane domain (IM), and cytoplasmic domain, (semaphorin) 6A (MGDI:G1120327, GB|NM\_018744, evidence: BLASTN, 100%, match:182);"

/protein\_id="BAC30582.1"  
/db\_xref="G:2334985"  
/translation="MRPAALLICLLHLCAGAGFPDESEPISSHGNYTKQYVPVFGH  
KPCNTQRRHIDQIMKIMNRTJXVAARDHITYVDIDSHTEIISCKKLTKLQSG  
DVTCKMGKQKDECHNTKYLKKANDLIPVCOTNAPNSGNSRYVLTETGDSF  
GMAYPQDAKHANALEADGLYSATVDLFAIDAVIYSLGDSQRLRTVHKHSSKWLK  
FYPVQAGVQGYVALEFRAVENYTMCKFVFRVAQVCKNDMGSGQRYLEKQWTF  
LKARLLNSPVGSHSFLYNIILQAQTVJIRENGRDVLAITSTPTSPGSAVAYAMD  
IANVFTRCKEYQKSDVSTWTPVDPDERPKPGCGCAASSLEKATYSSPFDDITNL  
KTHPLMDAEAVPSIINRPLRTMYRILTKLAANAAGNYN"VYF="GSEKGIILK  
LARISSGFLNGLIEEMNYNPKQSYDQVEKIRKTMQMDRSLGSAVAFSTCVI  
KVPGLSGKNGKNTKTLASRQPCGWWYFESSCAHLSPLSRCTFEQDIERGNTDGLG  
DHNSEVAIRGKKNTVTLASRQPCGWWYFESSCAHLSPLSRCTFEQDIERGNTDGLG

BASE COUNT	905 a	768 c	745 g	912 t
ORIGIN	DOHNSFVAEJNGKDFVIVTISR			

Query Match 46.58; score 1437.6; 53 11; Length 3329;

Best Local Similarity 89.18 Fred. No. 0  
Matches 1551; Conservative 0; Mismatches 129; Pairs 0; Caps 0

Qy	Db
1	673

[illegible][illegible][illegible]

241 ATAGACACATCACACACACGGGAGAAATTATCTGTATCAAAAACCTGCACATGGAAATCAGA 300

Dd	9.3	ATAGACATATCCACACAGAGAAATTA	ACTGTATGCAAAAACTGACATGGAAATCTPAGA	972
Qy	301	CAGGCCGATGTAGACATGCAGAAATGAAGG	AAAAACATAAAGGATGAGTGGCACAACATTT	360
Dd	973	CAGGCTGACGTAGACATGCAGGATGAGGGG	AAAAACAATGAAGGATGAATGTGCACAATTC	1032
Qy	361	ATTAAGTCTTCTTAAGAAAAACGATGATGCA	TGTTTGTCTGTGGAACATAATGCCTTC	420
Dd	1033	ATTAAGTCTTCTTAAGAAAAACGATGATGCA	TGTTTGTCTGTGGAACATAATGCCTTC	1092
Qy	421	AACCTTCTCTGCAGAAACTATAAGATGGATAC	ATGGAACCAATTCGGCGATGAATTCAGC	480
Dd	1093	AACCTTCTCTGCAGAAACTATAAGATGGATAC	ATGGAACCAATTCGGCGATGAATTCAGC	1152
Qy	481	GGATGGCCAGATGCCATATGATGGCAACATG	CGCAAGTGGTGCACATGTTGCGATGGA	540
Dd	1153	GGATGGCCAGATGCCATATGATGGCAACATG	CGCAAGTGGTGCACATGTTGCGATGGA	1212
Qy	541	AAACTATATCAGCAACAGTACTGACTTCCTTG	CGCCATGACGCGAGTCATTTACCGGAGT	600
Dd	1213	AAACTATATCAGCAACAGTACTGACTTCCTTG	CGCCATGACGCGAGTCATTTACCGGAGT	1272
Qy	601	CTTGAGAAAGCCCTATCCCTGAGACCGTCA	GCAGAGATCAAAATGGTTGAAAGAACCA	660
Dd	1273	CTTGAGAAAGCCCTATCCCTGAGACCGTCA	GCAGAGATCAAAATGGTTGAAAGAACCA	1332
Qy	661	TACTTTGTTTAAAGCGTGGATATGAGAGAT	ATATATCTCTCTTCAGGGAATATGCA	720
Dd	1333	TACTTTGTTTAAAGCGTGGATATGAGAGAT	ATATATCTCTCTTCAGGGAATATGCA	1392
Qy	721	GTGGAGTATACACCATGGCAAGGAGTACTTT	CCCCAAGAGTGGCTCAGGTTGTAAAGAT	780
Dd	1393	GTAGAAATACAACATATGGGGAAGTTGTTT	CCCTAGGGTGGCTCAGTCTGTAAAGAT	1452
Qy	781	GATATGGGAGATCTCAAGAGTCCITGGACA	AAACAGTGGAGCTGCTTCTGAGCGCGC	840
Dd	1453	GACATGGGAGGCTCTCAAGAGTCCITGGACA	AAACAGTGGAGCTGCTTCTGAGCGCTGCG	1512
Qy	841	TTGAACTCTCAGTTCCTGGAGACTCTCAT	TTTTTATTTCAACATTCCTCAGCGATTCAC	900
Dd	1513	CTGAACCTGCTGGTGGCTGGAGACTCTCAT	TTTTTATTTCAACATTCCTCAGCGATTCAC	1572
Qy	901	GAITGATTCTGATATCAAGGGCGTGATGTG	TCTCTGGCAACGTTTCTACACCTTATAC	960
Dd	1573	GAITGATTCTGATATCAAGGGCGTGATGTG	TCTCTGGCAACGTTTCTACACCTTATAC	1632
Qy	961	AGCATCCCTGGTCTGCAGTCTGTGCCATG	CATGCTTGAACATTCCTCAGCGATTTTACT	1020
Dd	1633	AGCATCCCGAGTCTGCAGTCTGTGCCATG	CATGCTTGAACATTCCTCAGCGATTTTACT	1692
Qy	1021	GGGAGATTCAAGGAACAGAGTCTCTGTAT	TCACCTGGACACCAAGTCCCTGATGAAGA	1080
Dd	1693	GGGAGATTCAAGGAACAGAGTCTCTGTAT	TCACCTGGACACCAAGTCCCTGATGAAGA	1752
Qy	1081	GTCTCTAAGCCAGGCGAGGTTGTGTCTG	TGGCTCATCTCTCTTACAAGATATGCAACC	1140
Dd	1753	GTCTCTAAGCCAGGCGAGGTTGTGTCTG	TGGCTCATCTCTCTTACAAGATATGCAACC	1812
Qy	1141	TCCAAATGAGTTCCTTGATATACCTGAACT	CTCATCAGACGACCGCTCATGATGATGAG	1200
Dd	1813	TCCAAATGAGTTCCTTGATATACCTGAACT	CTCATCAGACGACCGCTCATGATGATGAG	1872
Qy	1201	GCAGTGCCCTCCATCTTCAACAGGCCATG	TGTCTCTCAGAACAAATGGTATGATACGCC	1260
Dd	1873	GCAGTGCCCTCCATCTTCAACAGGCCATG	TGTCTCTCAGAACAAATGGTATGATACGCC	1932
Qy	1261	ACCAAAATTCAGTGGACGACAGCTGCTG	GGGCCATATCAGAAATCACACTGTGTTTTCTG	1320
Dd	1933	ACCAAAATTCAGTGGACGACAGCTGCTG	GGGCCATATCAGAAATCACACTGTGTTTTCTG	1992
Qy	1321	GGATCAGAAAGGAATCACTGAGTCTCTG	CCAGGATCTTGGCAGGATAGGAAGCAGTTCCTA	2052

```

QY 1381 AATGACAGCCCTTTCTGAGGAGAGATGAGTGTATACAACTGTGAAATGCGAGGTATGAT 1410
Db 2053 AATGGCAGCCTTTCTGAGGAGAGATGAGTGTATACAACTGTGAAATGCGAGGTATGAT 2112
QY 1441 SGAGTCGAAGACAAAGAGCATGCGATGAGCTGGACAGAGCAAGCAGCTCTCTGTAT 2530
Db 2113 GGTGTAGAGACAAAGAGCATGCGATGAGCTGGACAGAGCAAGCAGCTCTCTGTAT 2172
QY 1501 GTTGCCCTTCTACCTGTGTGATAAGGTTCOCCTTGGCCGTGTGAAAGCATGCGCAAG 1560
Db 2173 GTTGCCCTTCTACCTGTGTGATAAGGTTCOCCTTGGCCGTGTGAAAGCATGCGCAAG 2232
QY 1561 TGTAAATAACCTGTATTGCTCCAGAGACCATATGTTGATGATGATGATGATGATGATGAT 1620
Db 2233 TGTAAATAACCTGTATTGCTCCAGAGACCATATGTTGATGATGATGATGATGATGATGAT 2292
QY 1621 GCTGTGAGCCATTTATCACCCAAACAGAGACTGACTTTTIGACAGAGACATAGAGCTGGC 1680
Db 2293 TCTGTGCCCTCTGTCACTTCTAGCAGACTGACATTTGAGCAGGACATGAGAGCTGGC 2352
QY 1681 AATACAGATGCTGTGGGAGAGTGTGACAAATTCCTTTGTGGCACGAATGGCATTCAGT 1740
Db 2353 AATACGAGCGCTAGAGAGTGTGACAAATTCCTTGTGGCACGAATGATGATGAGGACAT 2412

RESULT 4
CB605722
LOCUS 2411 bp mRNA linear F01 16-MAY-2003
DEFINITION ANGNNUC:MRPE3-00075-D4-WY placenta embryo D17 (10379) Rattus
norvegicus cDNA clone mrpe3-00075-d4, mRNA sequence.
ACCESSION CB605722
VERSION 1
KEYWORDS Rattus norvegicus
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 2411)
AUTHORS Angen EST Program.
TITLE Angen Fat EST Program
JOURNAL Unpublished
COMMENT Contact: Dan Fitzpatrick
Argen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 30675 row: d column: 4.
Location/Organism
1. 2411
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="mrpe3-00075-d4"
/lisue_type="placenta embryo"
/clone_lib="placenta embryo D17 (10379)"
/note="Vector: pSPOR1; placenta embryo D17"
BASE COUNT 655 a 589 c 557 g 505 t
ORIGIN
Query Match 34.5%; Score 1065.8; DB 14; Length 2411;
Best Local Similarity 79.5%; Pred. No. 2.1e-270;
Matches 1398; Conservative 0; Mismatches 147; Indels 214; Gaps 3;

QY 1338 CATCTTGAAGTTTGGCCAGAAATAGAGTGTCTTAATGACAGCTTTTCCT 1397
Db 11 CACTTGAAGTTTGGCCAGAAATAGAGTGTCTTAATGACAGCTTTTCCT 70
QY 1398 GAGGAGATGAGTGTATCAAGCTGTAAATGCAAGCTATGATGAGTTCGAGGACAAAG 1457
Db 71 GAGGAGATGAGTGTATCAAGCTGTAAATGCAAGCTATGATGAGTTCGAGGACAAAG 130
QY 1458 GATCATGGGCGATGCACTGGACAGAGCAAGCAGCTCTCTGTATGTGCTTCTACCTG 1517

```

```

Db 132 GATCATGGGCGATGCACTGGACAGAGCAAGCTGTCTCACTTTATGTTGCAATCTCTACTG 190
QY 1518 TGTGATAAAGGTTCCCTTGGCCGGTGTGAACACATGGGAAGTGTAAATAACCTGTAT 1577
Db 291 TGTGATAAAGGTTCCCTTGGCCGGTGTGAACACATGGGAAGTGTAAATAACCTGTAT 250
QY 1578 TGTCTCCAGAGACCCATATTTGGATGGATGAAGGAAGTGTGCTCGAGCCATTTATC 1637
Db 251 CGCTCCAGAGACCCATATTTGGTGGTGGTAAAGGAAGTGTCTCTGCCCCATCTGTC 310
QY 1638 ACCACAGCAGCAGACTGACATTTGAGCAGGACATAGAGCGTGGCAATACAGATGCTGCG 1697
Db 311 ACCCTTACAGCAGACTGGCTTTTGGCAAGACATAGAGCGTGGCAATACAGAGCGCTAG 370
QY 1698 GGAATGTCAAAATTCCTTTGTGGCAGCTGAATGGCATTCCAGTTCCCTTCTGCCCCAG 1757
Db 371 AGACTGTCAAAATTCCTTTGTGGCAGCTGAAT 401
QY 1758 AACCACATCAGATTCGACGCTCAAGAGGGGTATGATGTAGGGGAGGAATGCTGCACTG 1817
Db 402 402
QY 1818 GAAGCATCTGTTGACATCACCIGACAGCAGAGACCCCTTTGGGGGCGAGTGTCTTCCCAT 1877
Db 402 402
QY 1878 TCACCAAGACAGAGAGGAGTGTATGTTGGGAAAGTACCTCAAAAGGCCAGCAGCAGCTGT 1937
Db 402 402
QY 1938 TCCGTCACCCCTTGGCCATTCGATTCCTTGGCTTTGCTCAITGGGGCCCTCTCTCTC 1997
Db 446 TCCCTGACCCCTTTGGCCATTCGATTCCTTGGCTTTGCTCAITGGGGCCCTCTCTCTC 505
QY 1998 GGGCATCACCTCTACTGCTCTGTATCAITGGGCGCAAGAGCTGGCTGTGGTGCAGCG 2057
Db 506 CGCATCACTGCTATGTTGTGTGTGATCAGCGCGCAAGAGCTGGCGTAGTGCAGCG 565
QY 2058 CAAGGAGAGAGCTCACCCACTCGGCGCGGGCTCATGAGCAGCGGTACCAAGCTCAG 2117
Db 566 CAAGGAGAGAGCTCACCCACTCGGCGCGGGCTCATGAGCAGCGGTACCAAGCTCAG 625
QY 2118 CGGCTCTTTGGGAGAGCTCAATCCCAAGAGCCCAAGCGCGAGCCATCTCACGCCACT 2177
Db 625 TGGCTCTTTGGGAGAGCTCAATCCCAAGAGCCCAAGCGCGAGCCATCTCACGCCACT 685
QY 2178 CATGCAACAGCGCAAGCTCGCCACTCTCGGCAACAGCGCGCAAGCTCATTAAGCAGA 2237
Db 686 CATGCAACAGCGCAAGCTCGCCACTCTCGGCAACAGCGCGCAAGCTCATTAAGCAGA 745
QY 2238 CGAGCAGCAGCTGAGCTTACCGGCTCTCGGCGCGGGCTCAAGCCAGAGTCAAGCCAGCGT 2297
Db 746 CGAGCAGCAGCTTACCGGCTCTCGGCGCGGGCTCAAGCCAGAGTCAAGCCAGCGT 805
QY 2298 GAAGCGGAAGCCAGCGCGCGCGAGCTGGGAGAGAAACAGCACTCAATCAATG 2357
Db 806 GAAGCGGAAGCCAGCGCGCGAGCTGGGAGAGAAACAGCACTCAATCAATG 865
QY 2358 CTGACAAAGAGCAGTACCGCCCTCGGCTCTCGGCTCTCGGCTCTCGGCTCTCGGCT 2417
Db 866 CTGACAAAGAGCAGT 880
QY 2418 GGCTTCCCCAGCCATCTCCAGCGTGTGGTGTCTGCTTCCCATCAGCCAGCGGCTACCA 2477
Db 881 -GGCTTCCCCAGCCATCTCCAGCGTGTGGTGTCTGCTTCCCATCAGCCAGCGGCTACCA 939
QY 2478 GCATGAGTACGTGACAGCCAGCCCAAAATGAGCGA---GGTGGCCAGATGCGCTGAGGA 2534
Db 940 GCAGGATACGTAGTACGCCCAAAATGAGCGAGTGGTGGCTCAGATGGCTGAGGAGGA 999
QY 2535 CCAGCGCGCCACACTGGAGTATAGACCATCAAGAGCAATCTCAGCAGCAAGAGTCCCCAA 2594
Db 2535 CCAGCGCGCCACACTGGAGTATAGACCATCAAGAGCAATCTCAGCAGCAAGAGTCCCCAA

```

1000 CCAGGCTCCACCCCTGGAGTAAAGACCAACAAAGAGCAGCTCAGCAGCAAGAGTCCCAA 1059  
Db  
2595 CCATGGGCTGAGACTGTGGAGAACCTGGAGAGCTCCGCCCAAGCTTCCAGAGCGGA 2654  
QY  
1060 CCATGGGCTGAGACTGTGGAGAACCTGGAGAGCTCCGCCCAAGCTTCCAGAGCTGA 1119  
Db  
2655 GGCCTCCCTGAGTCCCGGGAGCCTCCCTGTCTCAGACCGGTGTAAGCAAGCGGTGGA 2714  
QY  
1120 GGCCTCCCTGAGTCCCGGGAGCCTCCCTGTCTCAGACCGGTGTAAGCAAGCGGTGGA 1179  
Db  
2715 AATGACACACTCTCTCTCCAGGGGTGACTAAGAGAGCTACCCACAGCACTCGCT 2774  
QY  
1260 AATGACACACTCTCTCTCCAGGGGTGACTAAGAGAGCTACCCACAGCACTCGCT 1325  
Db  
2775 CAGGAGAGCCAGCAGGCAACATCTCTCAGAGCAACAGCAACAGCAACAGCAACAGCA 2834  
QY  
1240 CAGGAGAGCCAGCAGGCAACATCTCTCAGAGCAACAGCAACAGCAACAGCAACAGCA 1299  
Db  
2835 CAGGAGAGCCAGCAGGCAACATCTCTCAGAGCAACAGCAACAGCAACAGCAACAGCA 2894  
QY  
1300 CAGGAGAGCCAGCAGGCAACATCTCTCAGAGCAACAGCAACAGCAACAGCAACAGCA 1359  
Db  
2895 GGTGAGTCTATCAGGTCAGAGCAACATCTCTCAGAGCAACAGCAACAGCAACAGCAAC 2954  
QY  
1360 GGTGAGTCTATCAGGTCAGAGCAACATCTCTCAGAGCAACAGCAACAGCAACAGCAAC 1419  
Db  
2955 GGTGAGTCTATCAGGTCAGAGCAACATCTCTCAGAGCAACAGCAACAGCAACAGCAAC 3014  
QY  
1420 GGTGAGTCTATCAGGTCAGAGCAACATCTCTCAGAGCAACAGCAACAGCAACAGCAAC 1479  
Db  
3015 GGTGAGTCTATCAGGTCAGAGCAACATCTCTCAGAGCAACAGCAACAGCAACAGCAAC 3074  
QY  
1480 GGTGAGTCTATCAGGTCAGAGCAACATCTCTCAGAGCAACAGCAACAGCAACAGCAAC 1539  
Db  
3075 CAATGATCGGTGATACAA 3093  
QY  
1540 CAATGATCGGTGATACAA 3558  
Db

RESULT 5  
BM450002  
LOCUS  
DEFINITION  
AGENCORT\_639382 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5528003  
5', mRNA sequence.  
ACCESSION  
BM450002  
VERSION  
BM450002.1 GI:18495042  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

REFERENCE  
1. (bases 1 to 1030)  
Mammalia: Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: rgs@bgl.nhl.nih.gov  
Tissue Procurement: ATCC/NCBI/DTIP  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM12204 row: c column: 12  
High quality sequence stop: 733.  
Location/Qualifiers  
1. .1030  
/organism="Homo sapiens"  
/mol\_type="cDNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5528003"  
/tissue\_type="melanotic melanoma"

FEATURES  
source

/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_72"  
/note="Organ: Skin; Vector: pCMV-Sport6; Site: 1; Note:  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.  
Average insert size 2 kb. Library constructed by Life  
Technologies."  
BASE COUNT 268 a 236 c 257 g 266 t 3 others  
ORIGIN  
Query Match 29.1%; Score 899.4; DH 12; Length 1030;  
Host Local Similarity 97.4%; Pred. No. 1.5e-226;  
Matches 956; Conservative 0; Mismatches 19; Indels 7; Gaps 4;  
QY 767 GGAGGATCTCAAGAGTCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 846  
Db 1 GGAGGATCTCAAGAGTCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 60  
QY 847 TGCACAGTCTCTGAGAGTCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 906  
Db 63 TGCACAGTCTCTGAGAGTCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 120  
QY 907 ATTCTATTAACGAGGCGTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 966  
Db 121 ATTCTATTAACGAGGCGTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 180  
QY 967 CCGTGGTCTGAGAGTCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 1026  
Db 181 CCGTGGTCTGAGAGTCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 240  
QY 1027 TTCAAGGAGCAAGTCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 1086  
Db 241 TTCAAGGAGCAAGTCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 300  
QY 1087 AACCCAGGCGGAGTCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 1146  
Db 301 AACCCAGGCGGAGTCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 360  
QY 1147 GAGTCCCTGATGATACCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 1206  
Db 361 GAGTCCCTGATGATACCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 420  
QY 1207 CCTCCATCTTCAGAGCGGATCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 1266  
Db 421 CCTCCATCTTCAGAGCGGATCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 480  
QY 1267 ATTGAGTGGAGACAGTCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 1326  
Db 481 ATTGAGTGGAGACAGTCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 540  
QY 1327 GAGAGGAGATCATCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 1386  
Db 541 GAGAGGAGATCATCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 600  
QY 1387 AGCCTTTCTCTGAGAGTCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 1446  
Db 601 AGCCTTTCTCTGAGAGTCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 660  
QY 1447 GAAGACAAAGGATCATCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 1506  
Db 661 GAAGACAAAGGATCATCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 720  
QY 1507 TTTCTACCTGTGTATTAAGGTTCCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 1566  
Db 721 TTTCTACCTGTGTATTAAGGTTCCTCTGAGAACAGTGGAGTGGTTCCTGAAAGGCGGCTTGAAC 780  
QY 1566 AAAAAGCTGATTTGCTCCAGAGAGCCATTTGAGTGGATGAGTAAAGAGGTTGGGCTG 1625  
Db 781 AAAAAGCTGATTTGCTCCAGAGAGCCATTTGAGTGGATGAGTAAAGAGGTTGGGCTG 840  
QY 1626 CAGCAGTTTATCAGCCACAGCAGACTGCTTTGAGCAGGACATAGAGCGTGGCAGTAC 1685  
Db 841 CAGCAGTTTATCAGCCACAGCAGACTGCTTTGAGCAGGACATAGAGCGTGGCAGTAC 900



Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp

HRI Human cDNA project: 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

Location/Qualifiers

1. 868  
/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="PLACE200407"

/tissue\_type="placenta"

/clone\_lib="p.ACE2"

/note="vector: pHE-RSFL3"

230 a 197 c 206 g 236 t 5 others

BASE COUNT

ORIGIN

Query Match 25.8%; Score 826.2; Db 9; Length 868;

Best Local Similarity 98.3%; Pred. No. 9.7e-258;

Matches 855; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

QY 368 TTCCTCTAAGAAACAGATGATGATGTTGTCTCTGGAACTCAATGCGCTCAAGCCCT 427

DB 1 TTCTCTAAGAAACAGATGATGATGTTGTCTCTGGAACTCAATGCGCTCAAGCCCT 40

QY 428 CCTCAGAACTATAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 487

DB 6 CTTGAGAACTATAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 120

QY 488 CCAGTCCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 547

DB 121 CCAGTCCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 280

QY 548 ACTCAGCCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 607

DB 181 ACTCAGCCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240

QY 608 AAAGCCCTACCTCGGACCTCAAGCAGATTCAAATGCTTGAAGAACCACTTCTG 657

DB 241 AAAGCCCTACCTCGGACCTCAAGCAGATTCAAATGCTTGAAGAACCACTTCTG 380

QY 568 TTCAAGCCGTGGATTACGGAGATTAATCTTCTCTTTCAGGGAATACAGATGGAT 727

DB 301 TTCAAGCCGTGGATTACGGAGATTAATCTTCTCTTTCAGGGAATACAGATGGAT 360

QY 728 ATACACCATGGGAAGTACTTTTCCCAAGTGGCTCAGTTTGTGAAGATGATGG 787

DB 361 ATACACCATGGGAAGTACTTTTCCCAAGTGGCTCAGTTTGTGAAGATGATGG 420

QY 788 GAGGATCTCAAGAGTCTCTGAGAAACAGTGGAGTCTTCTGAAAGCGGTGGAAC 847

DB 421 GAGGATCTCAAGAGTCTCTGAGAAACAGTGGAGTCTTCTGAAAGCGGTGGAAC 480

QY 848 GCTCAGTCTCTGAGAGTCTCTGATTTTATTCACATTCACAGGAGTACACATGTA 907

DB 481 GCTCAGTCTCTGAGAGTCTCTGATTTTATTCACATTCACAGGAGTACACATGTA 540

QY 908 TTGCTATCAACGGGGGTGATGTTGCTCGGCAAGTTTCTACACCTTATAACAGCATCC 967

DB 541 TTGCTATCAACGGGGGTGATGTTGCTCGGCAAGTTTCTACACCTTATAACAGCATCC 600

QY 968 CTGGGTCTCAGTCTGTGCTATGACATGTGACATGTGACATGTGACATGTGACAT 1027

DB 601 CTGGGTCTCAGTCTGTGCTATGACATGTGACATGTGACATGTGACATGTGACAT 660

QY 1028 TCAAGAACAGAGTCTCTGATTCACATTCACAGGAGTCTCTGATGATGAGAGTCTCTTA 1087

DB 661 TCAAGAACAGAGTCTCTGATTCACATTCACAGGAGTCTCTGATGATGAGAGTCTCTTA 720

QY 1088 AGCCAGGCCAGGTGCTGTCTGCTCCTATCCTCTTGAAGATATGCAACCTCAATG 1147

DB 721 AGCCAGGCCAGGTGCTGTCTGCTCCTATCCTCTTGAAGATATGCAACCTCAATG 780

QY 1148 ACTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1207

DB 781 ACTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 848

QY 1208 CTTGATCTTCAAGAGGCAATGTTCTGTA 1237

DB 839 CTTGATCTTCAAGAGGCAATGTTCTGTA 868

RESULT 8

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

EMAIL

ISSUE

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA

CDNA



```

Db      188 GAAAGATATGCAACCTCCCAATGAGTTCCTGTGATGATACCTGAAATTCATGAGAGCGAC 247
QY      1185 CGGCTCATGATGAGCAGGCGGCTCCATCTTCACACAGCCCAATGGTTCGTGAGAGCAATG 1245
Db      245 CGGCTCATGATGAGCAGGCGGCTCCATCTTCACACAGCCCAATGGTTCGTGAGAGCAATG 307
QY      1245 GTTAGATACGGGCTTACCAAAATTCGAGTGGACAGAGGCTGCTGGGCCAATCATCAATCAG 1305
Db      308 GTCAGATACGGGCTTACCAAAATTCGAGTGGACAGAGGCTGCTGGGCCAATCATCAATCAG 367
QY      1305 ACTCTGCTTTCTCGATCAGAGAGGAATCACTTCATCTTTCAGATTTTGGCCAGATAGCA 1365
Db      368 ACTGTGGTTCCTGGATCAGAGAGGAAATCATCTTGAATTTTGGCCAGATAGCA 427
QY      1366 AATAGTGGTTCCTAAATGACAGCCTTTTCCTGGAGGATGAGTGTTCACAACTCTGAA 1425
Db      428 AATAGTGGTTCCTAAATGACAGCCTTTTCCTGGAGGATGAGTGTTCACAACTCTGAA 487
QY      1426 AATGACGATATGATGAGTTCGAGACAAAGAGATCATGGGATGAGCTGGACAGACGA 1485
Db      488 AATGACGATATGATGAGTTCGAGACAAAGAGATCATGGGATGAGCTGGACAGACGA 547
QY      1486 AGCAGTCTCTGATGCTGGGCTCTACCTGTGTATATAAGTTCCTGCGGGTGT 1545
Db      548 AGCAGTCTCTGATGCTGGGCTCTACCTGTGTATATAAGTTCCTGCGGGTGT 607
QY      1545 GAACGACATGGGAATGTMAAANAAGTGTATATGCTTCAGAGACCATATTTGTGATGG 1605
Db      608 GAACGACATGGGAATGTMAAANAAGTGTATATGCTTCAGAGACCATATTTGTGATGG 667
QY      1606 ATAAAGAGAGGTGCTGCTGACGCAATTTATACCCACAGACAGACTGACTGACGAC 1665
Db      668 ATAAAGAGAGGTGCTGCTGACGCAATTTATACCCACAGACAGACTGACTGACGAC 727
QY      1665 GACATAGAGCGGGCAATACAGATGGTTCGGGACTGTGCAATTTTGTGACAT 1724
Db      728 GACATAGAGCGGTGGCAATACAGATGGTTCGGGACTGTGCAATTTTGTGACAT 787
QY      1725 GAATGGCATTCAGTTCCTGCTTCGACAGACATACCCAGATCAGATTCGACAGATCAANA 1784
Db      788 GAATGGCATTCAGTTCCTGCTTCGACAGACATACCCAGATCAGATTCGACAGATCAANA 846
QY      1785 GGGGTATGATGCTATGAGGAGGANTGCTGAGTGTGAATCAT 1824
Db      847 AGGATATGAGTCTCTATAGAGAGCAATATCTGAGCTGAGAGAT 890

Bg769297 887 bp mRNA linear EST 15-MAY-2001
602742838F1 NIH_MGC_45 Homo sapiens cDNA clone IMAGE:4872704 5',
mRNA sequence.
Bg769297
Bg769297.1 GI:14079950
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
2 (bases 1 to 887)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: gqabbs@mail.nih.gov
Tissue Procurement: AIRC/DCTD/BTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

```

```

Plate: LUCM1749 row: k column: 09
High quality sequence stop: 862.

FEATURES
    source
        1..887
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /cclone="IMAGE:4872704"
            /tissue_type="melanotic melanoma, high MDR (cell line)"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_49"
            /note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGACAGAG(G). Size-selected >500bp for average insert size
            1.8kb. Library constructed by Ling Hong in the laboratory
            of Gerald M. Rubin (University of California, Berkeley)
            using ZAP-cDNA synthesis kit (Stratagene) and Superscript
            II RT (Life Technologies). Note: this is a NIH_MGC
            library."
BASE COUNT 206 a 300 c 243 g 138 t
ORIGIN
    Query Match 25.48; Score 784.4; DB 12; Length 887;
    Best Local Similarity 95.88; Pred. No. 4e-196;
    Matches 849; Conservative 0; Mismatches 31; Indels 6; Gaps 4;

QY 1903 CGGAAAGATTACCTCAAGAGCCAGACAGCTGGTTCCTCCCTCACCCTCTTGGCAITGCA 1962
Db 2 CGGAAAGATTACCTCAAGAGCCAGACAGCTGGTTCCTCCCTCACCCTCTTGGCAITGCA 61
QY 1963 GTCATCTGGCTTCTCATGCTGGGGCGCTCTTCGGGGCATCACCCTCTACTGCTGTG 2022
Db 62 GTCATCTGGCTTCTCATGCTGGGGCGCTCTTCGGGGCATCACCCTCTACTGCTGTG 121
QY 2023 GATCATCGGTCGAAGAGCGGTGGTGGTGCAGCGCAAGGAGAGGTCACCCACTG 2082
Db 122 GATCATCGGTCGAAGAGCGGTGGTGGTGCAGCGCAAGGAGAGGAGCTCACCCTG 181
QY 2083 CGCGGGGCTTCATGACAGCGGTGACCAAGCTCAGCGGCTCTTTGGGGACACTCAATC 2142
Db 182 CGCGGGGCTTCATGACAGCGGTGACCAAGCTCAGCGGCTCTTTGGGGACACTCAATC 241
QY 2143 AAGACCCCAAGCGCGGAGGCTCTCTCAGCCACTCATGTCAGACGCGACAGCTGGCACT 2202
Db 242 AAGACCCCAAGCGCGGAGGCTCTCTCAGCCACTCATGTCAGACGCGACAGCTGGCACT 301
QY 2203 CGCGGCAACAGCGGCAAGAAGTCTCTTAAAGCAGACAGCAGCAGCAGCCGACGCG 2262
Db 302 CGCGGCAACAGCGGCAAGAAGTCTCTTAAAGCAGACAGCAGCAGCAGCCGACGCG 361
QY 2263 CTCCCGCAGCCAGAGTCAACCCCAAGCTGACAGAGCGGAGGAGCCAGCCGCGGAGC 2322
Db 362 CTCCCGCAGCCAGAGTCAACCCCAAGCTGACAGAGCGGAGGAGCCAGCCGCGGAGC 421
QY 2323 CGCGAGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2382
Db 422 CGCGAGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 481
QY 2383 GGCTCCCTGTGATTCGACGAGGAGTCTGCGGCTGCGGCGCTCCGCGAGCAGATCCGAGC 2442
Db 482 GGCTCCCTGTGATTCGACGAGGAGTCTGCGGCTGCGGCGCTCCGCGAGCAGATCCGAGC 541
QY 2443 GTGCTGTCTCTGCTCCATCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2502
Db 542 GTGCTGTCTCTGCTCCATCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 601
QY 2503 ATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2561
Db 602 ATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 661
QY 2562 CATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2620

```

```

Db      662 CATCAAGAACATCTCAAGCAAGAGTCCCAACCATGGGGTGAACCTGTGAGAAC 721
Qy      2621 TGGACAGCTGCCGCCCAAGATTTCACAGCGGGAGGCTCCCTGGTCCCGCGGA--GC 2678
Db      722 TGGACAGCTGCCGCCCAAGATTTCACAGCGGGAGGCTCCCTGGTCCCGGTAGCC 781
Qy      2679 CTCCTGTCTCAGACGGGTCTAAGCAAGCGCTGGAAATGACCAACACCTCTCTCTACGG 2718
Db      782 TCCCTGTCTTACAGCGGTCTAAGCAAGCGCTGGAAATGACCAACCTCTCTCTCTACGG 841
Qy      2739 GCTT--GACTATAAGAGAGCTACCCACAGCACTGCTCTACAGAA 2782
Db      842 GGTAGACTTATAGAGCACTTAACACAGCAATGCTTCAAGACA 887

RESULT 10
AL543344
LOCUS      1231 bp mRNA linear EST 31-MAY-2003
DEFINITION AL543344 Homo sapiens PLACENTA cOT 25-NORMALIZED Homo sapiens cDNA
clone CSOD1001YL17 5'-PRIME, mRNA sequence.
ACCESSION AL543344
VERSION    AL543344.2 GI:31255191
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Placentalia; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE  1 (bases 1 to 1201)
AUTHORS   Li, W.B., Gruber, C., Jessup, J., and Pelley, D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished
COMMENT   On Feb 15, 2001 this sequence version replaced gi12875822.
            Genoscope
            Contact: Genoscope
            BP 191 9306 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by life technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 2864.r for
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CSOD1001CF090PL&cl=ster-2864.r. Contact :
            Peng Liang Email : liliang@lifetech.com URL :
            http://fulllength.invitrogen.com/Invitrogen Corporation 1600
            Paraday Avenue Genoscope sequence ID : CSOD1001CF090PL.
FEATURES
            Location/Qualifiers
                1..1201
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CSOD1001YL17"
                /tissue_type="PLACENTA cOT 25-NORMALIZED"
                /clone_lib="Homo sapiens PLACENTA cOT 25-oligo(dT)
                /note="1st strand cDNA was primed with a NotI-oligo(dT)
                primer. Five prime end enriched, double-strand cDNA was
                digested with NotI and cloned into the NotI and EcoRV
                sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 277 a 373 c 315 g 214 t
ORIGIN
Query Match 25.3% Score 783.4; DB 9; Length 1201;
Best Local Similarity 96.4%; Pred No. 8.8e-196;
Matches 838; Conservative 5; Mismatches 2; Indels 5; Gaps 4;
Qy 1714 TTGTGGGCTAATGAGGATTCACATTCCTCTTTGGGAGCAACACCAATCAGATTCG 1773
Db 221 TTGTGATGAAGAGAGGATTCACATTCCTCTTTGGGAGCAACACCAATCAGATTCG 280
Qy 1774 AGGGCTCAAGAGGGATTCACATTCCTCTTTGGGAGCAACACCAATCAGATTCG 1833
Db 281 AGGGCTCAAGAGGGATTCACATTCCTCTTTGGGAGCAACACCAATCAGATTCG 349
Qy 1834 TCACCTGACAGCAGACCTTTGGGAGCAACACCAATCAGATTCGATTCACAGAGAG 1893

```

```

Db      341 TCACCTGACAGCAGACACCCCTTTGGGGGAGTGTCTTCCCAATATCAACAAGCAAG 400
Qy      1894 GGAGTGATTTCGGGAAAGTTACTTCAAGAGCCAGCAGCTGGTTCCGCTCACCTCTTG 1953
Db      401 GGAGTGATTTCGGGAAAGTTACTTCAAGAGCCAGCAGCTGGTTCCGCTCACCTCTTG 460
Qy      1954 GCATTTGAGTATCTCTGCTTTTCATGAGGGGGCTTCTTCTCGGGATCATCCGCTTAC 2013
Db      461 GCATTTGAGTATCTCTGCTTTTCATGAGGGGGCTTCTTCTCGGGATCATCCGCTTAC 520
Qy      2014 TCGGCTGTGATCATCTGCGGAGAGAGCTGTGTGTGTCAGGCAAGGAGAGAGAGCTC 2073
Db      521 TCGGCTGTGATCATCTGCGGAGAGAGCTGTGTGTGTCAGGCAAGGAGAGAGAGCTC 580
Qy      2074 ACCCACTGCGGCGGGGTCTCATGAGCAAGCTGTCACCAAGCTCAGCGGGCTCTTTGGGAG 2133
Db      581 ACCCACTGCGGCGGGGTCTCATGAGCAAGCTGTCACCAAGCTCAGCGGGCTCTTTGGGAG 640
Qy      2134 ACTCAATCCAAAGAGCCAAAGCGGAGCGCATCTGTCAGCGCCACTCATGCACAACGGCAAG 2193
Db      641 ACTCAATCCAAAGAGCCAAAGCGGAGG--CATPCTTCAGCGCCACTCATGCACAACGGCAAG 699
Qy      2194 CTGCGCACTTCCCGGCAACACAGCGGCAAGATGCTCATTAAGCAGACACCAACCTGAC 2253
Db      700 CTGCGCACTTCCCGGCAACACAGCGGCAAGATGCTCATTAAGCAGACACCAACCTGAC 759
Qy      2254 CTGAGGGCTCTCCACACCGGAGAGTCAACCCAGAGCTGTCAGCAGAGAGCGGAGCCAGC 2313
Db      760 CTGAGGGCTCTCCACACCGGAGAGTCAACCCAGAGCTGTCAGCAGAGAGCGGAGCCAGC 819
Qy      2314 CCGGCGAGCGCGGATTTGGGAGAGAGACCAAGACCTCATCAATGGCTGCACAAAGAGCATG 2373
Db      820 CCGGCGAGCGCGGATTTGGGAGAGAGACCAAGACCTCATCAATGGCTGCACAAAGAGCATG 879
Qy      2374 CCGCCCATGGCTTCCCTGTGATTCACAGGACCTTGGCGGCTTCCGCGGCTTCCGCGGAGC 2433
Db      880 CCGCCCATGGCTTCCCTGTGATTCACAGGACCTTGGCGGCTTCCGCGGCTTCCGCGGAGC 939
Qy      2434 ATCCCGAGGCTGTGTCTGCTGCTCATCAGCAGAGGCTTACGACATGAGTGTGAC 2493
Db      940 ATCCCGAGGCTGTGTCTGCTGCTCATCAGCAGAGGCTTACGACATGAGTGTGAC 993
Qy      2494 CAGCCCAATAGGAGGATGCGCCAGATGCGCGCTGTGGAGGAGCGGCGCCACACTGGAG 2553
Db      1000 --AGCCCAATAGGAGGATGCG--CCAGATGGCGCTGGAGGCCAGG--CGSCACACTGGAG 1055
Qy      2554 TATAGACCATCAGAGACATCTCAGCAG 2582
Db      1056 TATAGACCATCAGAGAAATTTAGCAGCAG 1084

```

```

RESULT 11
B0678536
LOCUS      848 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_8208159 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6262572
5', mRNA sequence.
VERSION    B0678536
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE  1 (bases 1 to 848)
AUTHORS   NIH-MGC http://mgc.nhl.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT   Contact: Robert Srausberg, Ph.D.
            Email: csabps@mail.nih.gov
            Tissue Procurement: DCTD/PTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: L1CM2425 row: b column: 13  
High quality sequence stop: 595.  
Location/Qualifiers

FEATURES  
source

1. 848  
/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MAG3:6262572"

/tissue\_type="melanotic melanoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_112"

/note="Organ: skin; Vector: pCMT7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT 210 a 207 c 231 g 199 t 1 others

Query Match 25.28; Score 779.4; DB 13; Length 848;

Best Local Similarity 98.68; Pred. No. 8; e-1.95;

Matches 830; Conservative 0; Mismatches 12; Indels 4; Gaps 4;

QY 1286 CTGGGCGATATCAGATCAGTCTGGTCTTTCTGGGATCAGAGGGAATCATCTTCA 1345

DB 1 CTGGGCGATATCAGATCAGTCTGGTCTTTCTGGGATCAGAGGGAATCATCTTCA 50

QY 1346 AGTTTCTGGGCGAATAGGAATAGTGGTTTCTTAATGACAGCCTTTCTCTGGAGGAA 1405

DB 61 AGTTTCTGGGCGAATAGGAATAGTGGTTTCTTAATGACAGCCTTTCTCTGGAGGAA 120

QY 1406 TGAGTGTTTACACTCTGAPAAATGAGGTATGATGCTGCAAGCAAAAGATCATGG 1465

DB 121 TGAGTGTTTACACTCTGAPAAATGAGGTATGATGCTGCAAGCAAAAGATCATGG 160

QY 1466 GCATGAGCTGGGACAGAGCAGTCTCTCTATGTTGGCTCTGCTCTGCTGTGTGATAA 1525

DB 181 GCATGAGCTGGGACAGAGCAGTCTCTCTATGTTGGCTCTGCTCTGCTGTGTGATAA 240

QY 1526 AGGTTCCTCTGGGCGGTGTGACGACATGGGAATGTAAATAAAGCTGTATGCTTCA 1585

DB 241 AGGTTCCTCTGGGCGGTGTGACGACATGGGAATGTAAATAAAGCTGTATGCTTCA 300

QY 1586 GAGACCATATTTGGATGATTAAGGAAGTCTGCTGCTGCAAGCCATTATACCCCAACA 1645

DB 301 GAGACCATATTTGGATGATTAAGGAAGTCTGCTGCTGCAAGCCATTATACCCCAACA 360

QY 1646 GCAGCTGAGTTTTCAGCAGGATAGAGCGTGGCAATACAGATGCTCTGGGGAGTGTG 1705

DB 361 GCAGCTGAGTTTTCAGCAGGATAGAGCGTGGCAATACAGATGCTCTGGGGAGTGTG 420

QY 1706 ACAATTCCTTTGGGCACTGAATGGGCTATCCAGTCTCCCTCTGGGACGACAAACAA 1765

DB 421 ACAATTCCTTTGGGCACTGAATGGGCTATCCAGTCTCCCTCTGGGACGACAAACAA 480

QY 1766 CAGATTTCAGCGGTTCAGAGGGGTATGATCTTAGGGAGGGAATGCTTGGAGTGAAGCATC 1825

DB 481 CAGATTTCAGCGGTTCAGAGGGGTATGATCTTAGGGAGGGAATGCTTGGAGTGAAGCATC 540

QY 1826 TGCTTGACTACCTTCAGCAGCAGAGCCCTTTGGGGCAGTGTCTTCCCATATCAGCAAG 1885

DB 541 TGCTTGACTACCTTCAGCAGCAGAGCCCTTTGGGGCAGTGTCTTCCCATATCAGCAAG 600

QY 1886 ACAAGAGGGAGTCTATTCGGGAAAGTTACCTCAAGGCCACGACGAGCTGTTCCCGTCA 1945

DB 601 ACAAGAGGGAGTCTATTCGGGAAAGTTACCTCAAGGCCACGACGAGCTGTTCCCGTCA 660

QY 1946 CCCTCTTGGGCATTCAGCAATCTCTGCTCTTGGTATGGGGCCCGTCTTCTCGGCGATCA 2005

DB 661 CCCTCTTGGGCATTCAGCAATCTCTGCTCTTGGTATGGGGCCCGTCTTCTCGGCGATCA 720

QY 2006 CCCTCTACTGCTCTCTATCTGCTGCGGCAAGA-CGTGGCTGTGTGTGAG-CGCAAGA 2063

DB 721 CCCTCTACTGCTCTCTATCTGCTGCGGCAAGA-CGTGGCTGTGTGTGAG-CGCAAGA 780

QY 2064 GAAGGAGCTCACCACCT-CGCGCGGGGCTCCATGAG-CAGGCTCACCAGCTCAGCGC 2121

DB 781 AAGAGGCTCACCACCTCAGCGGGGGGCTCATGAGCCAGGCTCACCAGCTTCCCGG 840

QY 2122 CTTCTT 2127

DB 841 CCCTCT 846

RESULT 12

LOCUS BU172225

DEFINITION BU172225 883 bp mRNA linear EST 04-SEP-2002

5', mRNA sequence.

ACCESSION BU172225

VERSION BU172225.1 G1:22686209

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 883)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: rgs@bbs-rcmail.nih.gov

Tissue Procurement: DCTD/VRP

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrived by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1CM2347 row: m column: 15

High quality sequence stop: 606.

FEATURES

source

1. 883

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/tissue\_type="melanotic melanoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_112"

/note="Organ: skin; Vector: pCMT7; Site\_1: XhoI; Site\_2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGACGAG(G). Library constructed by Ling Hong in the

laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a

NIH\_MGC Library."

BASE COUNT 215 a 229 c 238 g 201 t

ORIGIN

Query Match 25.08; Score 772.4; DB 13; Length 883;

Best Local Similarity 98.68; Pred. No. 6e-193;

Matches 779; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1301 ATCAGCTGTGGTCTTCTGGGATCAGAGGGAATCATCTTGAAGCTTTTGGCCAGAA 1360

DB 14 ATCCAGCTGTGGTCTTCTGGGATCAGAGGGAATCATCTTGAAGCTTTTGGCCAGAA 73

QY 1361 TAGGAAA:AGTGGTGTTCCTTAATGACAGCCTTTTCC:GGAGGAGATGAGTGTTTACAACT 1420



```

QY 2316 CGCAGCGCGC-AGTGGGAGGAGACA-CAAGCTCATTAATCGGCGCAAGAGACATG 2373
DB 721 CGGCAACGCGGAGTGGGAGAGAACANNAAGCTCATCAATGCCCTGCACNAAGAGACATG 780
QY 2374 CCCCCCATGG--CTCCCTCTGTATTCACAGAGAGCTG-----CCCTGGGGCTCCCG 2427
DB 781 CCCCCCATGGGGCTCCCTGCTGATTTCCACGAGCGCTCCCTGCGGGGCTCCCGC 840
QY 2428 AGCCACATCCCGAGCGTCGG 2419
DB 841 AGCACATCCCGCCCAAGCGGG 862

RESULT 14
BUB38082
LOCUS BUB38082 874 bp mRNA linear EST 15-OCT-2002
DEFINITION AGENCOURT_8071571 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6088226
5' mRNA sequence.
ACCESSION BUB38082
VERSION BUB38082.1 GI:24022477
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Contact: Robert Strausberg, Ph.D.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Email: cyapbs-remail.nih.gov
Tissue Procurement: DfID/DFP
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2426 row: b column: 03
High quality sequence stop: 652.
Location/Qualifiers
1..874
/organism="Homo sap-ens"
/clone="IMAGE:608224"
/db_xref="taxon:9606"
/issue_type="melanoma, cell: line"
/lab_host="D410B (phage-resistant)"
/clone_lib="NIH_MGC_112"
/notes="Organ: Skin; Vector: pCTR7; Site: 2; XhoI: Site 2;
Note: cDNA made by oligo-dT priming. Directionally cloned
into EcoR/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 213 a 211 c 247 g 203 t
ORIGIN
Query Match
Best Local Similarity 99.2% Score 767.4; DB 13; Length 874;
Matches 771; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1286 CTGGGCGCATATCAGAAATCACTGGTTTCTGGGATCAGAGAGGAATCATCTTGA 1345
DB 1 CTGGGCGCATATCAGAAATCACTGGTTTCTGGGATCAGAGAGGAATCATCTTGA 60
QY 1346 AGTTTCTGGCGAGATAGAAATAGTGGTTTCTAAATAGACACCTTTCTCTGGAGAGA 1405
DB 61 AGTTTCTGGCGAGATAGAAATAGTGGTTTCTAAATAGACACCTTTCTCTGGAGAGA 120

```

```

QY 1406 TGAATGTTTAAACACTCTGAAAAATGCAGCTATGATGAGTGCAGCAAAAAGGATCATGG 1465
DB 121 TGAATGTTTAAACACTCTGAAAAATGCAGCTATGATGAGTGCAGCAAAAAGGATCATGG 180
QY 1466 GCATGCACTGGACAGAGCAAGCAGCTCTGTATGTTGCGTCTCTACCTGTTGATAA 1525
DB 181 GCATGCACTGGACAGAGCAAGCAGCTCTGTATGTTGCGTCTCTACCTGTTGATAA 240
QY 1526 AGTTTCCCTTGGCGGTGTGAAGCAGATGGAGTGTAAAAAACCTGTTTCCTTCCA 1585
DB 241 AGTTTCCCTTGGCGGTGTGAAGCAGATGGAGTGTAAAAAACCTGTTTCCTTCCA 309
QY 1586 GAGACCATATGTGGATGGATAAGGAAGTGTGCTGAGGCAATTTATCAACCAACA 1645
DB 301 GAGACCATATGTGGATGGATAAGGAAGTGTGCTGAGGCAATTTATCAACCAACA 360
QY 1646 CGAGACTCAGTTTGGACAGGACATAGAGCGTGGCAATACAGATGGTGTGGGAGCTGC 1705
DB 361 CGAGACTCAGTTTGGACAGGACATAGAGCGTGGCAATACAGATGGTGTGGGAGCTGC 420
QY 1706 ACAATTTCCTTTGTGSCACTGAATGGGCAATCCAGTTCCTTCCCGCAGCAACACCAT 1765
DB 421 ACAATTTCCTTTGTGSCACTGAATGGGCAATCCAGTTCCTTCCCGCAGCAACACCAT 480
QY 1766 CAGATTCGACGCCICAGAGGGGTATGAGTCTAGGGGAGGAATGCTGAGCTGGAGCATC 1825
DB 481 CAGATTCGACGCCICAGAGGGGTATGAGTCTAGGGGAGGAATGCTGAGCTGGAGCATC 540
QY 1826 TCGTTGACTTCACCTGACAGCAGACACCCCTTTGGGGGAGTGTCTTCCCATATCACCAG 1885
DB 541 TCGTTGACTTCACCTGACAGCAGACACCCCTTTGGGGGAGTGTCTTCCCATATCACCAG 600
QY 1886 ACAGAAAGGAGTGATTCGGGAAGTACCTCAAGGCCAGCAGCAGCTGTTCCCGCA 1945
DB 601 ACAGAAAGGAGTGATTCGGGAAGTACCTCAAGGCCAGCAGCAGCTGTTCCCGCA 660
QY 1946 CCTCTTGGCATTCAGTTCCTGCTTCTGCTATGCGGGGCGCTCTCTCGGGCATCA 2005
DB 661 CCTCTTGGCATTCAGTTCCTGCTTCTGCTATGCGGGGCGCTCTCTCGGGCATCA 720
QY 2006 CCTCTTACTGCTGTGTATCATCTCGGCCAAGAGCTGCTGTGTGTCAGGCGAAG 2062
DB 721 CCTCTTACTGCTGTGTATCATCTCGGCCAAGAGCTGCTGTGTGTCAGGCGAAG 777

RFSL115
BQ683009
LOCUS BQ683009.1 593 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_820RC14 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6262587
5' mRNA sequence.
ACCESSION BQ683009
VERSION BQ683009.1 GI:21795688
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 593)
AUTHORS NIH-MGC http://mgc.ncbi.nlm.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cyapbs-remail.nih.gov
Tissue Procurement: DfID/DFP
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2425 row: c column: 04
High quality sequence stop: 621.
Location/Qualifiers

```

FEATURES

```
source 1..890
/organism="Homo sapiens"
/genotype="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6262587"
/issue_type="melanotic melanoma, cell line"
/lab_host="BHI0B (phage-resistant)"
/clone_lib="NIH_MGC_112"
/notes="organ skin; Vector: pORF7; Size:1; XhoI: Site 2;
ECORI: cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GCCACGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."
```

```
BASE COUNT 222 a 220 c 245 g 204 t
ORIGIN
Query Match 24.6%; Score 760.6; E: 13; Length 890;
Best Local Similarity 97.0%; Pred. No. 46-146;
Matches 808; Conservative (-); Mismatches 19; Indels 6; Gaps 7;
QY 1286 CTGGGCGCATCAGATACATCTGTGTTTCTGTGGATCAGAGAGGGAATCATCTTGA 1345
Db 1 CTGGGCGCATCAGATACATCTGTGTTTCTGTGGATCAGAGAGGGAATCATCTTGA 60
QY 1346 AGTTTGGCCAGATAGGAATAGTGGCTTTCTAAAGACAGGCTTTCTGTGAGSACA 1405
Db 61 AGTTTGGCCAGATAGGAATAGTGGCTTTCTAAAGACAGGCTTTCTGTGAGSACA 120
QY 1406 TGAGTGTTCAGACTCGAAGATGAGGATATGATGAGTGTGAGAGGAGAGATCATGG 1465
Db 121 TGAGTGTTCAGACTCGAAGATGAGGATATGATGAGTGTGAGAGGAGAGATCATGG 180
QY 1466 GCATGACGTGGACAGAGGAGGCTCTGTGTATGTTGCTTCTACCTGTGTGATAA 1525
Db 181 GCATGACGTGGACAGAGGAGGCTCTGTGTATGTTGCTTCTACCTGTGTGATAA 240
QY 1526 AGTTCCCTTCGCGGCTGGACACATCGGAAAGTGTAAAGAACCTGTATGCTGCA 1585
Db 241 AGTTCCCTTCGCGGCTGGACACATCGGAAAGTGTAAAGAACCTGTATGCTGCA 300
QY 1586 GAGACCCATATTGTGGATGATAAAGGAGGTGGTGGCTGACGACATTTATCACCCACA 1645
Db 301 GAGACCCATATTGTGGATGATAAAGGAGGTGGTGGCTGACGACATTTATCACCCACA 360
QY 1646 GCAGACTGACTTTTGACAGGACATAGAGGCTGGCAATACAGATGGTCTGGGGACTGTC 1705
Db 361 GCAGACTGACTTTTGACAGGACATAGAGGCTGGCAATACAGATGGTCTGGGGACTGTC 420
QY 1706 ACAATTCCTTTTGGGCACTGAATGGGCAATCCAGTTCCCTCTGGCCAGCACACACAT 1765
Db 421 ACAATTCCTTTTGGGCACTGAATGGGCAATCCAGTTCCCTCTGGCCAGCACACACAT 480
QY 1766 CAGATTCGACGGCTCAAGAGCGGTATGATCTAGGGGAGGAAATGCTGGACTGGAGGATC 1825
Db 481 CAGATTCGACGGCTCAAGAGCGGTATGATCTAGGGGAGGAAATGCTGGACTGGAGGATC 540
QY 1826 TGCTTGACTCAGCTTGACAGACAGACCCCTTTGGGGGAGCTGTCTTCCCATATATCAGCAAG 1885
Db 541 TGCTTGACTCAGCTTGACAGACAGACCCCTTTGGGGGAGCTGTCTTCCCATATATCAGCAAG 600
QY 1886 ACAAGAGGGAGTGATTCGGGAAAGTTACCTCAAGGGCCAGGACAGCTGGTTCGGTCA 1945
Db 601 ACAAGAGGGAGTGATTCGGGAAAGTTACCTCAAGGGCCAGGACAGCTGGTTCGGTCA 660
QY 1946 CCCTCTTGGCCATTCGACATCGCTGGCTTCTGTCAAGGGGGGCTCTCTGGGGATCA 2005
Db 661 CCCTCTTGGCCATTCGACATCGCTGGCTTCTGTCAAGGGGGGCTCTCTCTGGGGATCA 720
QY 2006 CCGTCTACTGGCTGTGTATCATCGGGGCAAG-AGGTGGCTGTGGTGGAGGCGCAGG-- 2062
|||||
```

```
Search completed: September 30, 2003, 22:02:09
Job time : 5947.36 secs
```

1	3038	98.2	3495	11	US-09-951-053-3	Sequence 3, Appl
2	3036	98.2	3498	11	US-09-957-187-3	Sequence 3, Appl
3	3030.4	98.0	4250	11	US-09-957-187-84	Sequence 84, Appl
4	2698	87.2	3333	11	US-09-951-053-5	Sequence 5, Appl
5	2698	87.2	3333	11	US-09-957-187-5	Sequence 5, Appl
6	1881.6	60.8	1890	11	US-09-991-053-29	Sequence 29, App
7	1881.6	60.8	1890	11	US-09-957-187-29	Sequence 29, App
8	1881.6	60.8	1890	11	US-09-957-187-82	Sequence 82, App
9	604	19.5	845	10	US-09-933-381-920	Sequence 920, App
10	565	18.3	1472	9	US-09-925-299-60	Sequence 60, App
11	565	18.3	1472	11	US-09-925-299-60	Sequence 60, App
12	555.4	18.0	3524	12	US-10-391-413-3	Sequence 3, Appl
13	553.8	17.9	3721	11	US-09-931-836-34	Sequence 34, Appl
14	553.8	17.9	3721	12	US-10-035-977-34	Sequence 34, Appl
15	553.8	17.9	3721	12	US-10-137-870-543	Sequence 543, App
16	553.8	17.9	3721	12	US-10-140-018-543	Sequence 543, App







934	Db	GTGGAGTATAACACCATGGGAAAGGTAGATTTCCTCCCAAGASTGGCTCAGGTTTGTATAAGAT	993
781	QY	GATATGGGAGGATCTCAAGAGAGTCCTGGAGAACAGTGGACGTGCTTCTGGAGSGGGC	840
994	Db	GATATGGGAGGATCTCAAGAGAGTCCTGGAGAACAGTGGACGTGCTTCTGGAGSGGGC	1053
841	QY	TTGAACTGGTCACTTCTGGAGACTTCATTTTATTTCAAAATTCGACGACGATTAGA	900
1054	Db	TTGAACTGGTCACTTCTGGAGACTTCATTTTATTTCAAAATTCGACGACGATTAGA	1113
901	QY	GATGTANTGCTATCAAGGGGCTGAATGTCTCTGGCAAGGTTTCTACACTTATAC	960
1114	Db	GATGTANTGCTATCAAGGGGCTGAATGTCTCTGGCAAGGTTTCTACACTTATAC	1173
961	QY	AGCATCCCTGGGTCTGCAGTCTGTGCTATGACATGCTTGACATTCGCCAGTGTTTTACT	1020
1174	Db	AGCATCCCTGGGTCTGCAGTCTGTGCTATGACATGCTTGACATTCGCCAGTGTTTTACT	1233
1021	QY	GGGAGATTCAAGAAACAGAAATCTCTGTGATTCACCTGGACACCACTTCCTGATGACGA	1080
1234	Db	GGGAGATTCAAGAAACAGAAATCTCTGTGATTCACCTGGACACCACTTCCTGATGACGA	1293
1081	QY	GTTCCTAAAGCCAGGCGACGTTCTCTGTGGGTCTATCTGCTTAGAAAGATATGCAGC	1140
1294	Db	GTTCCTAAAGCCAGGCGACGTTCTCTGTGGGTCTATCTGCTTAGAAAGATATGCAGC	1353
1141	QY	TCCAAATGATTCCTGATGATACACCTGATGACTTCATCAAGACAGTCAAGGCTCATGATGAG	1200
1354	Db	TCCAAATGATTCCTGATGATACACCTGATGACTTCATCAAGACAGTCAAGGCTCATGATGAG	1413
1201	QY	GCAGTGGCTTCCATCTTCACAGGCCATGTTTCTGTGAGAACATGCTCATATACCCGCTT	1260
1414	Db	GCAGTGGCTTCCATCTTCACAGGCCATGTTTCTGTGAGAACATGCTCATATACCCGCTT	1473
1261	QY	ACCAAAATTCGAGTGGACACAGTCTGTGGGCAATATCAAAATCACACTGTGGTCTTTCTG	1320
1474	Db	ACCAAAATTCGAGTGGACACAGTCTGTGGGCAATATCAAAATCACACTGTGGTCTTTCTG	1533
1321	QY	GGATCAGAGGAAGGAATCATCTTGAAGTCTTTGGCCAGATACGAATAGTGTTTCTGA	1380
1534	Db	GGATCAGAGGAAGGAATCATCTTGAAGTCTTTGGCCAGATACGAATAGTGTTTCTGA	1593
1381	QY	AATGACAGGCTTTTCTGTGGAGGATGAGTGTGTTACACTCTGAAAAATGCAGCTATGAT	1440
1594	Db	AATGACAGGCTTTTCTGTGGAGGATGAGTGTGTTACACTCTGAAAAATGCAGCTATGAT	1653
1441	QY	GGAGTCGAGACAAAAAGATCATGGCATGCAGCTGGACAGACCAAGCTCTCTGAT	1500
1654	Db	GGAGTCGAGACAAAAAGATCATGGCATGCAGCTGGACAGACCAAGCTCTCTGAT	1713
1501	QY	GTTCGCTTCTCTACCTGTGTGATAAAGGTTCCGCTTGGGCGGTGTGAACGACATGGGA	1560
1714	Db	GTTCGCTTCTCTACCTGTGTGATAAAGGTTCCGCTTGGGCGGTGTGAACGACATGGGA	1773
1561	QY	TGTAAAAAACCTGTATTTGCCGCCAGACCCCATATTGTGGATGGATTAAGGAAGTGCT	1620
1774	Db	TGTAAAAAACCTGTATTTGCCGCCAGACCCCATATTGTGGATGGATTAAGGAAGTGCT	1833
1621	QY	GCTCGAGGCAATTTACAGCCAAACAGACAGACTGACTTTTGAAGCAGACATAGAGCGTGGC	1680
1834	Db	GCTCGAGGCAATTTACAGCCAAACAGACAGACTGACTTTTGAAGCAGACATAGAGCGTGGC	1893
1681	QY	AATACAGATGGTCTGGGGACTGTACAAATTCCTTTTGTGGACATGATGGCATTCGACT	1740
1894	Db	AATACAGATGGTCTGGGGACTGTACAAATTCCTTTTGTGGACATGATGGCATTCGACT	1953
1741	QY	TCCCTCTTCCCGCAGCACACCATCAGATTCCACGGCTCAAGAGGGGTATGACTCTAGG	1800
1954	Db	TCCCTCTTCCCGCAGCACACCATCAGATTCCACGGCTCAAGAGGGGTATGACTCTAGG	2013
1801	QY	GGAGGAATGCTGCAGTGGAGGCAATCTCCTTTGACTACCTGACAGCAGACACCTTTGGG	1860

Db	2014	GGAGGAATGGTGGAGCACTCTGCTTGACTCACTGACAGCAGACCTTTGGGG	2071
Qy	1861	GCAGTGTCTCCCATTAATCAACCAAGACAAGAGGGAGTGAATCGGGAAAGTTCACCTCAAA	1920
Db	2074	GCAGTGTCTCCCATTAATCACCAAGACACAGAGGGAGTGATTCGGAAAGTACCTCAAA	2133
Qy	1921	GGCCAGCAGCAGTGGTTCGGGICACDCTCTTTGGCCATTGCACTCATCTCTGGTTCGTGTC	1980
Db	2134	GGCCAGCAGCAGTGGTTCGGCTCACCTCTTGCCATTGCACTCATCTCTGGCTTTCTGTC	2193
Qy	1981	ATGGGGGCGCTCTCTCGSGCATCACGCTACTGGTCTGTGATCNTGGCGCAAGAC	2040
Db	2194	ATGGGGGCGCTCTCTCGGGCATCACGGTACTGCGGCTGTGTGATCATCGGCCAAAGAC	2253
Qy	2041	GTGGCTGTGGAGGAGAGGAGCTCACCCACTTCGCGCGGGGCTCCATGAGC	2100
Db	2254	GTGGCTGTGGAGGAGAGGAGGAGCTCACCCACTTCGCGCGGGGCTCCATGAGC	2313
Qy	2101	AGGCTCACCAAGCTCAGCGGCTCTTTGGGACACTCAATCCAAAGACCCAAAGCGGGAG	2160
Db	2314	AGGCTCACCAAGCTCAGCGGCTCTTTGGGGACACTCAATCCAAAGACCCAAAGCGGGAG	2373
Qy	2161	GGCATCTCACGCGCACTCATGCGACAACGGCAAGCTGCCACTCCCGGCAACCGGCAAG	2220
Db	2374	GGCATCTCACGCGCACTCATGCGACAACGGCAAGCTGCCACTCCCGGCAACCGGCAAG	2433
Qy	2221	ATGCTTATTAAAGCAGCAGCAGCAGCTCTGAGCTTGACGGGCGCTCCGCCACCCAGAGTCA	2280
Db	2434	ATGCTTATTAAAGCAGCTAGCAGCAGCTTGAGGCTTGACGGGCTCCGCCACCCAGAGTCA	2493
Qy	2281	ACCCCAACGCTGAGCAGAGAGGAGAGCCAGCGGGGAGCGGCAACCGGAGTGGAGAGAGC	2340
Db	2494	ACCCCAACGCTGAGCAGAGAGGAGAGCCAGCGGGGAGCGGCAACCGGAGTGGAGAGAGC	2553
Qy	2341	CAGAACCTCATTAATGCTTCACAAGAGACATGGCCCCCAATGGGCTCTGTGATTCCG	2400
Db	2554	CAGAACCTCATTAATGCTTCACAAGAGACATGGCCCCCAATGGGCTCTGTGATTCCG	2613
Qy	2401	ACGGACTGCGCTGTGGGCTTCCCCAGCCACATCCCCAGCTGTGTCTCTGCCCATC	2460
Db	2614	ACGGACTGCGCTGTGGGCTTCCCCAGCCACATCCCCAGCTGTGTCTCTGCCCATC	2673
Qy	2461	ACGAGCAGGGCTACCAAGCATGAGTAGTGGGACAGCCCAAAATGAGCAGAGTGGCCCCAG	2520
Db	2674	ACGAGCAGGGCTACCAAGCATGAGTAGTGGGACAGCCCAAAATGAGCAGAGTGGCCCCAG	2733
Qy	2521	ATGGCGTGTGAGGACAGGCGGCCACACTGGAGTATAAGACCATCAAGGAACATCTCAGC	2580
Db	2734	ATGGCGTGTGAGGACAGGCGGCCACACTGGAGTATAAGACCATCAAGGAACATCTCAGC	2793
Qy	2581	AGCAAGAGTCCCAACCATGGGGTGAACCTTTGTGGAGAACCTGGACAGCTGCCCCCAAA	2640
Db	2794	AGCAAGAGTCCCAACCATGGGGTGAACCTTTGTGGAGAACCTGGACAGCTGCCCCCAAA	2853
Qy	2641	GTTCACAGCGGGAGGGCTCCCTGGGTCCCGCGGAGCCCTCCCTGTCTCAGACCGGTCTA	2700
Db	2854	GTTCACAGCGGGAGGGCTCCCTGGGTCCCGCGGAGCCCTCCCTGTCTCAGACCGGTCTA	2913
Qy	2701	AGCAAGCGGCTGGAAATATGACCACTCTCTTCCTACGGGGTTGACTATAGAAGAGCTAC	2760
Db	2914	AGCAAGCGGCTGGAAATATGACCACTCTCTTCCTACGGGGTTGACTATAGAAGAGCTAC	2973
Qy	2761	CCACAGAACTCGCTCAGCAGAGGCCACACGCGCCACC-ACTCTCAAAAGAAACAACTAA	2819
Db	2974	CCACAGAACTCGCTCAGCAGAGGCCACCTGACCACCTACTCTCTCATCAAGAGCAACTAA	3053
Qy	2820	CTCTCTCAATTCCTCTC-ACCTCTCCAGAAACCAAGAGCTTTGGCAGGGAGAGCAACCCGC	2878
Db	3054	CTCTCTCAATTCCTCTC-ACCTCTCCAGAAACCAAGAGCTTTGGCAGGGAGAGCAACCCGC	3093
Qy	2879	CGCCGCGCCGCGAGAGGTTGACTTCATCCAGGTTGCACAGCTCCCAAGCATCTGGCCAGG	2938
Db	3094	CGCCGCGCCGCGAGAGGTTGACTTCATCCAGGTTGCACAGCTCCCAAGCATCTGGCCAGG	3153

QY	2939	CCGTGACATGTTCTTAGAGGCACGCCAGACTCAACACCTATAAACCATTATGAACAAGTAAAGGGT	2998
Db			
Db	3154	CCGTGACTGTCCTGAGACACCGCAGGCTCAGGCTGACAGCTGACAACTGACAAAGCTGACAAAGCT	3213
QY	2939	TGAAGCGTACCGCTTGCTTAAGCGCGGAGTGTCAGCGCCCAAAACCATCTCTTTGCTCCCGCTT	3056
Db			
Db	3214	TGAGCGTACAGCGCTGCTTAAGCGCGGAGTGTCAGCGCCCAAAACCATCTCTTTGCTCCCGCTT	3273
QY	3059	CACATCATGATGAAGCCCAAATGATCGGTGTACATAA	3093
Db			
Db	3274	CCACATCATGATGAAGCCCAAATGATCGGTGTACATAA	3308

RESULT 3  
US-09-957-187-84  
Sequence 84, Application US/09957187  
Publication No. US20030054514A;  
GENERAL INFORMATION:  
APPLICANT: Shimkets, Richard A.  
APPLICANT: LaRoche, William  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY  
FILE REFERENCE: 1S966-540 CIP  
CURRENT APPLICATION NUMBER: US/09/957,187  
CURRENT FILING DATE: 2000-09-19  
PRIOR APPLICATION NUMBER: 60/123,667  
PRIOR FILING DATE: 1999-03-09  
PRIOR APPLICATION NUMBER: 09/520,781  
PRIOR FILING DATE: 2003-03-03  
PRIOR APPLICATION NUMBER: 60/234,082  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: 63/233,798  
PRIOR FILING DATE: 2000-09-19  
PRIOR APPLICATION NUMBER: 60/174,485  
PRIOR FILING DATE: 2000-01-04  
NUMBER OF SEQ ID NOS: 85  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 84  
LENGTH: 4250  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (250)...(3940)  
US-09-957-187-84

Query Match 98.0%; Score 3030.4; DS 11; Length 4250;  
Best Local Similarity 98.3%; Pred. No. 6;  
Matches 3092; Conservative 0; Mismatches 1; Indels 51; Gaps 17

QY	1	ATGAGCTCAGAAGCGTTGCTGCTATATTCACACTGCTACACTTGGCTGGGCGCTGTTTC	60
Db	250	ATGAGCTCAGAAGCGTTGCTGCTATATTCACACTGCTACACTTGGCTGGGCGCTGTTTC	309
QY	61	CCAGAAGATTCGAGCCCATAGATTTGCGATGCGACATATACMAAACAGHATCCCGTG	120
Db	310	CCAGAAGATTCGAGCCCATAGATTTGCGATGCGACATATACMAAACAGATATCCCGTG	369
QY	121	TTTGTGGGCGCACAGCGAGGCGGACAGACAGACAGACAGGCGTGCACATCCAGATG	180
Db	370	TTTGTGGGCGCACAGCGAGGCGGACAGACAGACAGGCGTGCACATCCAGATG	429
QY	181	ATTATGATCATGACGAGCGCTGCTACATGCTGCTGCTAGCGAGCATATTTATCTGCTGAT	240
Db	430	ATTATGATCATGACGAGCGCTGCTACATGCTGCTGCTAGCGAGCATATTTATCTGCTGAT	489
QY	241	ATAGACAGCTCACAACGCGAGGAATTTNTGTAGCAAAAATCTGACATGGAAATCTAGA	300
Db	490	ATAGACAGCTCACAACGCGAGGAATTTNTGTAGCAAAAATCTGACATGGAAATCTAGA	549
QY	301	CAGGCCGATGTAGACACATGACGAATCAAGGGAAAAACATTAAGGATGAGTGCCACAACTT	360
Db	550	CAGGCCGATGTAGACACATGACGAATCAAGGGAAAAACATTAAGGATGAGTGCCACAACTT	609



Query Match				37.2% Score 2698; DB 11; Length 3333;			
Best Local Similarity				53.9% Pred. No. 0;			
Matches 2907; Conservative				(- Mismatches 21; Indels 167; Gaps 3;			
QY	1	ATGAGGTCAGAAAGCTTGGTCTGCTATATATACACATGCTACACTTGGCTGGGCTGGTTTC	60	DB	214	ATGAGGTCAGAAAGCTTGGTCTGCTATATATACACTTGGCTGGGCTGGTTTC	273
QY	61	CCAGAGGATTCGAGCCAGCAGCAATTTGGATGCGACTATACAAACATGATCCGGTG	120	DB	274	CCAGAGGATTCGAGCCAGCAGCAATTTGGATGCGACTATACAAACATGATCCGGTG	333
QY	121	TTTGTGGGCGACAGCCAGCAGCAACACACAGAGGACAGGCTGGACATCCAGATG	180	DB	334	TTTGTGGGCGACAGCCAGCAGCAACACACAGAGGACAGGCTGGACATCCAGATG	393
QY	181	ATTATGATCATGAAGCGGACCGCTCTACATTCCTGCTAGGAGCAATATATACCTTGAT	240	DB	394	ATTATGATCATGAAGCGGACCGCTCTACATTCCTGCTAGGAGCAATATATACCTTGAT	453
QY	241	ATAGACATCATCACACGGAAGAAATTTATTTGAGCAAAACACTGACATGGAAATCTAGA	300	DB	454	ATAGACATCATCACACGGAAGAAATTTATTTGAGCAAAACACTGACATGGAAATCTAGA	513
QY	301	CAGGCGGATGTAGACACATGAGAAATGAAGGAAACATTAAGATGAGTGCACACATTT	360	DB	514	CAGGCGGATGTAGACACATGAGAAATGAAGGAAACATTAAGATGAGTGCACACATTT	573
QY	361	ATTAAAGTCTCTTAAAGAAACCATGATGCTATGTTTGTCTGGAACATTAATGCTTC	420	DB	574	ATTAAAGTCTCTTAAAGAAACCATGATGCTATGTTTGTCTGGAACATTAATGCTTC	633
QY	421	AACCTTCTGAGAAATTAAGATGATACATTTGAGCAACCATTCGGGATGATTAACG	480	DB	634	AACCTTCTGAGAAATTAAGATGATACATTTGAGCAACCATTCGGGATGATTAACG	593
QY	481	GAAATGGCGAGTGGCCATATGATCCCAACATGCAACGTTGGACTGTTTCCAGATGGA	540	DB	594	GAAATGGCGAGTGGCCATATGATCCCAACATGCAACGTTGGACTGTTTCCAGATGGA	600
QY	541	AACTATCTACACACAGTCTACTGACTTCTTGGCAATGACGAGTCAATACAGTCAAT	600	DB	600	AACTATCTACACACAGTCTACTGACTTCTTGGCAATGACGAGTCAATACAGTCAAT	600
QY	601	CTTGAGAGAGGCTTACCTGGGAGCGCTGAGACAGCTTCAATATGCTTGAAGAGCA	660	DB	660	CTTGAGAGAGGCTTACCTGGGAGCGCTGAGACAGCTTCAATATGCTTGAAGAGCA	660
QY	661	TACTTGTTCAGGCGTGGATACGAGATATATATCTACTTCTTCTTCAAGGAAATAGCA	720	DB	720	TACTTGTTCAGGCGTGGATACGAGATATATATCTACTTCTTCTTCAAGGAAATAGCA	720
QY	721	GTGGAGTATACAGCATGGGAAAGTATTTCCACAGAGTGGCTCAGCTTNTAAGAAAT	780	DB	780	GTGGAGTATACAGCATGGGAAAGTATTTCCACAGAGTGGCTCAGCTTNTAAGAAAT	780
QY	781	GATATGGAGGATCTCAAGAGGCTCTGGAGAAACAGAGGAGTCTCTTCAAGGCGGCG	840	DB	840	GATATGGAGGATCTCAAGAGGCTCTGGAGAAACAGAGGAGTCTCTTCAAGGCGGCG	840
QY	841	TTTAACTGCTGAGTCTGAGGAGCTCTCATTTTATTTCAACATTTCTCAGGAGTATACA	900	DB	900	TTTAACTGCTGAGTCTGAGGAGCTCTCATTTTATTTCAACATTTCTCAGGAGTATACA	900
QY	901	GATGATGCTGATCAAGCGGCTGATGTTGGCTGGCAAGCTTTTACACCTTATTAAC	960	DB	960	GATGATGCTGATCAAGCGGCTGATGTTGGCTGGCAAGCTTTTACACCTTATTAAC	960
QY	961	AGCATCCCTGGGCTGCACTGCTGCTATGATGATGCTGATGCTGCTGCTGCTGCTGCT	1020	DB	1020	AGCATCCCTGGGCTGCACTGCTGCTATGATGATGCTGATGCTGCTGCTGCTGCTGCT	1020
QY	1021	GGAGATTCAGAGAAAGAGTCTCTCTGATTCACCTGACACACACACACACACACAC	1080	DB	1080	GGAGATTCAGAGAAAGAGTCTCTCTGATTCACCTGACACACACACACACACACAC	1080
QY	1081	GTCTTAAAGCAAGGACAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1140	DB	1140	GTCTTAAAGCAAGGACAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1140
QY	1141	TCGAATGATTCCTGATGATACCTGAACTTCATCAGACAGCAGCAGCAGCAGCAGC	1200	DB	1200	TCGAATGATTCCTGATGATACCTGAACTTCATCAGACAGCAGCAGCAGCAGCAGC	1200
QY	1201	GGAGTGGGCTGATTCACAGGAGCAGCTTCTGCTGAGCAATAGGAAATAGTGGTTTCTA	1260	DB	1260	GGAGTGGGCTGATTCACAGGAGCAGCTTCTGCTGAGCAATAGGAAATAGTGGTTTCTA	1260
QY	1261	ACCAAAATTCGAGTGGACACAGCTGCTGGCCCATATCAGAAATCAGACTGCTGTTTCTG	1320	DB	1320	ACCAAAATTCGAGTGGACACAGCTGCTGGCCCATATCAGAAATCAGACTGCTGTTTCTG	1320
QY	1321	GGATCAGAGAGGAGTATCTTGAAGTCTTTGGCAGCAATAGGAAATAGTGGTTTCTA	1380	DB	1380	GGATCAGAGAGGAGTATCTTGAAGTCTTTGGCAGCAATAGGAAATAGTGGTTTCTA	1380
QY	1381	AATGACAGCTTCTCTGAGGAGATGAGTGTTCACACTCTGAAATATGACAGCTATGAT	1440	DB	1440	AATGACAGCTTCTCTGAGGAGATGAGTGTTCACACTCTGAAATATGACAGCTATGAT	1440
QY	1441	GGAGTGGGCTGATTCAGGAGTGGACATGAGCTGGACAGAGCAGCAGCAGCAGCAGC	1500	DB	1500	GGAGTGGGCTGATTCAGGAGTGGACATGAGCTGGACAGAGCAGCAGCAGCAGCAGC	1500
QY	1501	GTGCGTCTCTACCTGTGTGATTAAGTTCCTTGGCCGCTGTCAGGAGATGGGAG	1560	DB	1560	GTGCGTCTCTACCTGTGTGATTAAGTTCCTTGGCCGCTGTCAGGAGATGGGAG	1560
QY	1561	TGTAAGGAGTGGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG	1620	DB	1620	TGTAAGGAGTGGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG	1620
QY	1621	GGTCTGAGGATTCATCAGCAACAGCAGCTGCTTTGAGCAGCAGCAGCAGCAGCAGC	1680	DB	1680	GGTCTGAGGATTCATCAGCAACAGCAGCTGCTTTGAGCAGCAGCAGCAGCAGCAGC	1680
QY	1681	AATACAGATGGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG	1740	DB	1740	AATACAGATGGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG	1740
QY	1741	TCCTCTTGGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	1800	DB	1800	TCCTCTTGGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	1800
QY	1801	GGAGGATGCTGAGTGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1860	DB	1860	GGAGGATGCTGAGTGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1860
QY	1861	GCAGTCTCTTCCCATATCAGCAAGCAAGAGGAGTGGTGGGAGGAGTGGGAGGAG	1920	DB	1920	GCAGTCTCTTCCCATATCAGCAAGCAAGAGGAGTGGTGGGAGGAGTGGGAGGAG	1920
QY	1921	GGCAGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1980	DB	1980	GGCAGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1980
QY	1981	ATGAGGAGTGGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG	2040	DB	2040	ATGAGGAGTGGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG	2040
QY	2041	GTGGCTGCTGAGCAGCAGGAGGAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGG	2100	DB	2100	GTGGCTGCTGAGCAGCAGGAGGAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGG	2100
QY	2089	GTGGCTGCTGAGCAGCAGGAGGAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGG	2148	DB	2148	GTGGCTGCTGAGCAGCAGGAGGAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGG	2148
QY	2101	AGGTCAGCAGGCTGAGCGGCTCTTTGGGAGACTCAATCCAAAGACCCAAAGCGGAG	2160	DB	2160	AGGTCAGCAGGCTGAGCGGCTCTTTGGGAGACTCAATCCAAAGACCCAAAGCGGAG	2160

2149 AGCGCACCAAGCTCAGCGGCTCTCTGTGGGACACCAATCCAAAGACCCAAAGCCCGAG 2208  
2161 GCCATGCTCAGCGGCACTCATCAGCAAGCGGAAGTCCGCACTCCCGGCAACAGCGGCAAG 2220  
2209 CCCATCTCAGCGGCACTCATCAGCAAGCGGAAGTCCGCACTCCCGGCAACAGCGGCAAG 2266  
2221 ATGCTCATTAAGCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2260  
2269 ATGCTCATTAAGCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2328  
2281 ACCCAACGCTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2340  
2329 ACCCAACGCTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2388  
2341 CAGAACCTCATCATGCTGCAGAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2400  
2399 CAGAACCTCATCATGCTGCAGAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2446  
2401 ACGGACCTGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTG 2460  
2449 ACGGACCTGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTG 2508  
2461 ACGGACGAGGCTTACAGCACTGAGTACGCTGACAGCAGCAGCAGCAGCAGCAGCAGCAG 2520  
2509 ACGGACGAGGCTTACAGCACTGAGTACGCTGACAGCAGCAGCAGCAGCAGCAGCAGCAG 2568  
2521 ATGCGCTGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2580  
2569 ATGCGCTGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2628  
2581 AGCAAGAGTCCCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2640  
2629 AGCAAGAGTCCCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2688  
2641 GTTCCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2700  
2689 GTTCCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2748  
2701 AGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2760  
2749 AGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2808  
2761 CCCAGGAACTCGCTCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2819  
2809 CCCAGGAACTCGCTCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2868  
2820 CTCTCCCAATTCCTCTC-ACTCTCCAGAACCCAGAGCCTTGGCAGGGGAGACACCCCG 2878  
2869 CCCGACAAATTCAGCTGACTTCAAGGACCCAGAGCCTTGGCAGGGGAGACACCCCG 2928  
2879 CCCGACAAATTCAGCTGACTTCAAGGACCCAGAGCCTTGGCAGGGGAGACACCCCG 2938  
2929 CCCGACAAATTCAGCTGACTTCAAGGACCCAGAGCCTTGGCAGGGGAGACACCCCG 2988  
2939 CCCGACAAATTCAGCTGACTTCAAGGACCCAGAGCCTTGGCAGGGGAGACACCCCG 3048  
2989 CCCGACAAATTCAGCTGACTTCAAGGACCCAGAGCCTTGGCAGGGGAGACACCCCG 3098  
3049 TGAAGCGTACGCGCTCGCTAAAGCGGAGCAGTACCCCAACCACTCTTCTCTCCCTTTT 3108  
3059 CCACATCCATGAGCCCAATGATCGGTATACATAA 3093  
3109 CCACATCCATGAGCCCAATGATCGGTATACATAA 3143

RESULT 5  
US-09-957-187-5  
; Sequence 5, Application US/09957137  
; Publication No. US2003005414A1  
; GENERAL INFORMATION:

APPLICANT: Shimkels, Richard A.  
APPLICANT: LaRoche, William  
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND PROTEINS ENCODED THEREBY  
FILE REFERENCE: 15966-540 CIP  
CURRENT APPLICATION NUMBER: US/09/957,187  
CURRENT FILING DATE: 2000-09-19  
PRIOR APPLICATION NUMBER: 60/123,667  
PRIOR FILING DATE: 1999-03-09  
PRIOR APPLICATION NUMBER: 09/520,781  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: 60/234,082  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: 60/233,798  
PRIOR FILING DATE: 2000-09-19  
PRIOR APPLICATION NUMBER: 60/174,485  
PRIOR FILING DATE: 2000-01-04  
NUMBER OF SEQ ID NOS: 85  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 3333  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (214)..(2865)  
NAME/KEY: misc feature  
LOCATION: (2882)  
OTHER INFORMATION: an n may be any one of a or t or g or c  
US-09-957-187-5

Query Match 87.2%; Score 2698; DH 11; Length 3333;  
Best Local Similarity 93.9%; Pred. No. 0;  
Matches 2907; Conservative 0; Mismatches 21; Indels 167; Gaps 3;

QY 1 ATGAGTCAGAGGCTGCTCTATATTCACACTGCTACACTTGGCTGGGCTGGTTTC 60  
DB 214 ATGAGTCAGAGGCTGCTCTATATTCACACTGCTACACTTGGCTGGGCTGGTTTC 273  
QY 61 CCAGAAATTCGTAGCAATCAGTATTCGATGGCACTATACAAACAGTATCCGGTG 120  
DB 274 CCAGAAATTCGTAGCAATCAGTATTCGATGGCACTATACAAACAGTATCCGGTG 333  
QY 121 TTGTGGCCACAGCCAGGAGGAGCAACACACAGAGGAGGAGGAGGAGGAGGAGGAGG 180  
DB 334 TTGTGGCCACAGCCAGGAGGAGCAACACACAGAGGAGGAGGAGGAGGAGGAGGAGG 393  
QY 181 ATTATGATCATGAACGGAACCCCTGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
DB 394 ATTATGATCATGAACGGAACCCCTGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 453  
QY 241 ATAGACATCATACACAGGAGAGAAATTTATTTAGCAAAACATGACATGGAATCTAGA 300  
DB 454 ATAGACATCATACACAGGAGAGAAATTTATTTAGCAAAACATGACATGGAATCTAGA 513  
QY 301 CAGGCGGATGTAGACACATGAGAGTGAAGGCAAAACATGAGGATGAGTCCCAACATTT 360  
DB 514 CAGGCGGATGTAGACACATGAGAGTGAAGGCAAAACATGAGGATGAGTCCCAACATTT 573  
QY 361 ATTAAGTCTTCTTAAGAAACAGTATGATGATGTTTGTCTGTGGAACTATGCTTC 420  
DB 574 ATTAAGTCTTCTTAAGAAACAGTATGATGATGTTTGTCTGTGGAACTATGCTTC 633  
QY 421 AACCTTCTCAGAACACTATAAGATGATACATTTGGAACCATTCGGGGAGTATTCAGC 480  
DB 634 AACCTTCTCAGAACACTATAAGATGATACATTTGGAACCATTCGGGGAGTATTCAGC 693  
QY 481 GGAATGGCAGATGCCCATATGATGCCAAACATGCCAACGTTGCACCTTTTCAGATGGA 540  
DB 694 GGAATGGCAGATGCCCATATGATGCCAAACATGCCAACGTTGCACCTTTTCAGATGGA 753  
QY 541 AACCTTCTCAGAACACTATAAGATGATACATTTGGAACCATTCGGGGAGTATTCAGC 600  
DB 754 AACCTTCTCAGAACACTATAAGATGATACATTTGGAACCATTCGGGGAGTATTCAGC 813







QY 1428 TCGAGCTATGATGGAGTGGAGAGAGAAAGCAATATGAGGATGAGCTGAGCTGAGAGAGAGAG 1428  
DB 1431 TCGAGCTATGATGGAGTGGAGAGAGAAAGCAATATGAGGATGAGCTGAGCTGAGAGAGAGAG 1440  
QY 1455 AGCTCTCTGATGTTGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1456  
DB 1441 AGCTCTCTGATGTTGGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1500  
QY 1549 CGACATGGGAAGTGTAAAGAAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1549  
DB 1501 CGACATGGGAAGTGTAAAGAAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560  
QY 1609 AAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1568  
DB 1561 AAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1520  
QY 1669 ATAGAGCTGGCAATACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1728  
DB 1621 ATAGAGCTGGCAATACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1580  
QY 1729 GGGCATTCAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1788  
DB 1681 GGGCATTCAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
QY 1789 TATGAGTCTAGGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1848  
DB 1741 TATGAGTCTAGGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
QY 1849 GACCTTTGGGGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1908  
DB 1801 GACCTTTGGGGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860  
QY 1909 AGTACTCTCAAGGCCAGCAGCAGCTGG 1936  
DB 1861 AGTACTCTCAAGGCCAGCAGCAGCTGG 1886

## RESULT 7

US-09-957-187-29  
Sequence 29, Application US/09951187  
Publication No. US2003005414A1  
GENERAL INFORMATION:  
APPLICANT: Shimkets, Richard A.  
APPLICANT: LaRoche, William  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY  
FILE REFERENCE: 15466-545 CIP  
CURRENT APPLICATION NUMBER: US/09/957,187  
CURRENT FILING DATE: 2000-03-19  
PRIOR APPLICATION NUMBER: 60/123,667  
PRIOR FILING DATE: 1999-03-09  
PRIOR APPLICATION NUMBER: 09/520,781  
PRIOR FILING DATE: 2000-03-23  
PRIOR APPLICATION NUMBER: 60/234,082  
PRIOR FILING DATE: 2000-03-20  
PRIOR APPLICATION NUMBER: 60/233,796  
PRIOR FILING DATE: 2000-03-19  
PRIOR APPLICATION NUMBER: 60/174,485  
PRIOR FILING DATE: 2000-01-04  
NUMBER OF SEQ ID NOS: 65  
SOFTWARE: PatentLis Ver. 2.1  
SEQ ID NO 29  
LENGTH: 1890  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1890)  
US-09-957-187-29

Query Match: 60.8%; Score 1881.6; DB 1; Length 1890;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1884; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 49 GGGGATGGTTCCTCCAGAAAGATTCAGGCAATCAGTATTTCCGATGCGCACTATACAAA 108  
DB 1 GGAATCGGCTTCCTCCAGAAAGATTCAGGCAATCAGTATTTCCGATGCGCACTATACAAA 60  
QY 109 CAGTATCGGCTTCCTCCAGAAAGATTCAGGCAATCAGTATTTCCGATGCGCACTATACAAA 168  
DB 61 CAGTATCGGCTTCCTCCAGAAAGATTCAGGCAATCAGTATTTCCGATGCGCACTATACAAA 120  
QY 169 GACATCCAGATGATATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 228  
DB 121 GACAACAGATGATATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
QY 229 TATATGTTGATATATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 288  
DB 181 TATATGTTGATATATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
QY 289 TGGAAATCTAGACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 348  
DB 241 TGGAAATCTAGACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
QY 349 TGGCAAACTTTATTAAGATTCCTTAAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 408  
DB 301 TGGCAAACTTTATTAAGATTCCTTAAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 360  
QY 409 ACTAATGCTTCAACCCCTTCCTGCAAGAACTATAGATGATGATGATGATGATGATGATGATGATGAT 468  
DB 361 ACTAATGCTTCAACCCCTTCCTGCAAGAACTATAGATGATGATGATGATGATGATGATGATGATGAT 420  
QY 469 GATGAATTCAGCGGAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 528  
DB 421 GATGAATTCAGCGGAATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
QY 529 TTGCGAGATGGAAGATATATCTCAGCCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 588  
DB 481 TTGCGAGATGGAAGATATATCTCAGCCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
QY 589 ATTACCGGAGTCTTGAGAAAGCCCTACCTGCGGAGCGTCAAGCAGATTCAAAATGG 648  
DB 541 ATTACCGGAGTCTTGAGAAAGCCCTACCTGCGGAGCGTCAAGCAGATTCAAAATGG 600  
QY 649 TTGAAGAAGACCATCTTTTCAAGCCGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 708  
DB 601 TTGAAGAAGACCATCTTTTCAAGCCGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
QY 709 AGGGAATAGCAGTGGATATACACCATGCGGAAAGTAGTTTCCCAAGAGTGGCTCAG 768  
DB 661 AGGGAATAGCAGTGGATATACACCATGCGGAAAGTAGTTTCCCAAGAGTGGCTCAG 720  
QY 769 GTTTGATGAATGATATGGAGGATCTCAAGAGTCTTGAGAAAGCTGCGGAGCGTTC 828  
DB 721 GTTTGATGAATGATATGGAGGATCTCAAGAGTCTTGAGAAAGCTGCGGAGCGTTC 780  
QY 829 CTGAAGCCCGCTTGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888  
DB 781 CTGAAGCCCGCTTGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
QY 889 CAGGAGTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948  
DB 841 CAGGAGTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
QY 949 ACACCTTATAACAGATCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1008  
DB 901 ACACCTTATAACAGATCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 960  
QY 1009 AGTGTGTTTACTGGGAGATTCAGGAGACAGAGTCTGCTGATTCAGCTGGACACCATTT 1068  
DB 961 AGTGTGTTTACTGGGAGATTCAGGAGACAGAGTCTGCTGATTCAGCTGGACACCATTT 1020  
QY 1069 CTTGATGAAGAGTCTGCTGAAGCCAGGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1128  
DB 1021 CTTGATGAAGAGTCTGCTGAAGCCAGGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
QY 1129 AGATATGCAACCTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188



[illegible]

```

RESULT 9
US-09-833-381-920/c
? Sequence 920, Application: US/0983381
? Patent No. US20020132090A1
? GENERAL INFORMATION:
? APPLICANT: Robison, Keith E.
? TITLE OF INVENTION: No. US20020132090A1el Nucleic acid and Protein Homologs
? FILE REFERENCE: 5800-119
? CURRENT APPLICATION NUMBER: US/09/833.361
? CURRENT FILING DATE: 2001-04-11
? PRIOR APPLICATION NUMBER: 09/516,448
? PRIOR FILING DATE: 2000-02-29
? NUMBER OF SEQ ID NOS: 2055
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 920
? LENGTH: 846
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-833-381-920

```

Query Match 19.5%; Score 604; DB 10; Length 846;  
Best Local Similarity 100.0%; Pred. No. 1.3e-178;  
Matches 604; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Oy		1	ATGAGGTGAGAAGCCCTTGCTGCTATATTTCACACTGCCTACACTTGTGCTGGGGTGGTTTC	60
Db		604	ATGAGGTGAGAAGCCCTTGCTGCTATATTTCACACTGCCTACACTTGTGCTGGGGTGGTTTC	545
Oy		61	CCAGAAGATTCTGAGCCAATCAGTAATTCGCATGCGCAACTATACAAAACAGTAGTCGCGTG	120
Db		544	CCAGAAGATTCTGAGCCAATCAGTAATTCGCATGCGCAACTATACAAAACAGTAGTCGCGTG	485
Oy		121	TTTGTGGGCCACAAGCCAGGACGGAAACCACACAGAGGCGACAGGCTGACATCCAGATG	180
Db		484	TTTGTGGGCCACAAGCCAGGACGGAAACCACACAGAGGCGACAGGCTGACATCCAGATG	425
Oy		181	ATTATGATCATCAAGCGGRACCTCTACATTGCTGCTAGGAGCACCATATATTACTGTGAT	240
Db		424	ATTATGATCATCAAGCGGRACCTCTACATTGCTGCTAGGAGCACCATATATTACTGTGAT	305
Oy		241	ATAGACATCATCACACGGAAGAAATTTATTGTAGCAAAAACTGACATGGAAATCTAGA	300
Db		364	ATAGACATCATCACACGGAAGAAATTTATTGTAGCAAAAACTGACATGGAAATCTAGA	305
Oy		301	CAGGCGGATGTAGACACATGCAGATGAAGGGGAAAAACAATAGGATGAGTGCCCAACTTT	360
Db		304	CAGGCGGATGTAGACACATGCAGATGAAGGGGAAAAACAATAGGATGAGTGCCCAACTTT	245
Oy		361	ATTAAAGTTCCTTCTAAGSAAAAACGATCATGCATTGTGTGCTGTGGAACATAATGCCTTC	420
Db		244	ATTAAAGTTCCTTCTAAGSAAAAACGATCATGCATTGTGTGCTGTGGAACATAATGCCTTC	185
Oy		421	AAGCCTTCCTGGAGAACTATAAGATGSGATCATGTGGAACCATTCGSGGATGAATTCAGC	480
Db		184	AAGCCTTCCTGGAGAACTATAAGATGSGATCATGTGGAACCATTCGSGGATGAATTCAGC	125
Oy		481	GGAATGGCCAGATGGCCAATGATGCCAAACAATGCCACGTTGCACACTGTTTGAGATGGA	540
Db		124	GGAATGGCCAGATGGCCAATGATGCCAAACAATGCCACGTTGCACACTGTTTGAGATGGA	65
Oy		541	AAACTATACTACGCCACAGTGACTGACTTCCCTGCCATTGACGCAGTCATTACCGGAGT	600
Db		64	AAACATACTACGCCACAGTGACTGACTTCCCTGCCATTGACGCAGTCATTACCGGAGT	5
Oy		601	CTTG 604	
Db		4	CTTG 1	

RESULT 10  
US-C9-925-299-60

Sequence 60, Application US/09925299  
Patent No. US20020055627A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA102  
CURRENT APPLICATION NUMBER: US/09/925,299  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US03/05883  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 1556  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 60  
LENGTH: 1472  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (129)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc\_feature  
LOCATION: (130)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-299-60

Query Match 18.38: Score 565; DB 9: Length 1472;  
Best Local Similarity 99.5%; Pred. No. 3.4e-166;  
Matches 576; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY	2515	GCCAGATGGCGGTGGAGACACAGGCGGCCACACTGAGTATAGACCACTCAAGGAACAT	2574
DB	13	GCCAGATGGCGGTGGAGACACAGGCGGCCACACTGAGTATAGACCACTCAAGGAACAT	72
QY	2575	CTCAGCAGCAAGAGTCCCAACCATGGGTGAACCTTGAGGAACTGGACAGCCTGGCC	2634
DB	73	CTCAGCAGCAAGAGTCCCAACCATGGGTGAACCTTGAGGAACTGGACAGCCTNNCC	132
QY	2635	CCCAAGTTCACAGCGGAGGCGTCCCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT	2694
DB	133	CCCAAGTTCACAGCGGAGGCGTCCCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT	192
QY	2695	GGTCTAAGCAAGGCGTGGAAATGCACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2754
DB	193	GGTCTAAGCAAGGCGTGGAAATGCACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	252
QY	2755	AGTACCCCGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	2814
DB	253	AGTACCCCGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	311
QY	2815	ACTAATCTCTCCAAATTCCT	2874
DB	312	ACTAATCTCTCCAAATTCCT	371
QY	2875	CG	2934
DB	372	CG	431
QY	2935	CAGGCGGTGACTGTCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	2994
DB	432	CAGGCGGTGACTGTCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	491
QY	2995	GGGCTGAAGCGTACGGCTCTGCTAAGCGCGGAGTACCGGCTTTCGCTCC	3054
DB	492	GGGCTGAAGCGTACGGCTCTGCTAAGCGCGGAGTACCGGCTTTCGCTCC	551
QY	3055	CTTCCACATCCATGAAGCGGAGTACCGGCTTTCGCTCC	3093
DB	552	CTTCCACATCCATGAAGCGGAGTACCGGCTTTCGCTCC	590

RESULT 11

US-09-925-299-50  
Sequence 60, Application US/09925299  
Patent No. US2003030406.7A9  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA102  
CURRENT APPLICATION NUMBER: US/09/925,299  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05883  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 1556  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 60  
LENGTH: 1472  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (129)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc\_feature  
LOCATION: (130)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-299-60

Query Match 18.38: Score 565; DB 11: Length 1472;  
Best Local Similarity 99.5%; Pred. No. 3.4e-166;  
Matches 576; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY	2515	GCCAGATGGCGGTGGAGACACAGGCGGCCACACTGAGTATAGACCACTCAAGGAACAT	2574
DB	13	GCCAGATGGCGGTGGAGACACAGGCGGCCACACTGAGTATAGACCACTCAAGGAACAT	72
QY	2575	CTCAGCAGCAAGAGTCCCAACCATGGGTGAACCTTGAGGAACTGGACAGCCTGGCC	2634
DB	73	CTCAGCAGCAAGAGTCCCAACCATGGGTGAACCTTGAGGAACTGGACAGCCTNNCC	132
QY	2635	CCCAAGTTCACAGCGGAGGCGTCCCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT	2694
DB	133	CCCAAGTTCACAGCGGAGGCGTCCCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT	192
QY	2695	GGTCTAAGCAAGGCGTGGAAATGCACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2754
DB	193	GGTCTAAGCAAGGCGTGGAAATGCACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	252
QY	2755	AGTACCCCGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	2814
DB	253	AGTACCCCGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	311
QY	2815	ACTAATCTCTCCAAATTCCT	2874
DB	312	ACTAATCTCTCCAAATTCCT	371
QY	2875	CG	2934
DB	372	CG	431
QY	2935	CAGGCGGTGACTGTCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	2994
DB	432	CAGGCGGTGACTGTCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	491
QY	2995	GGGCTGAAGCGTACGGCTCTGCTAAGCGCGGAGTACCGGCTTTCGCTCC	3054
DB	492	GGGCTGAAGCGTACGGCTCTGCTAAGCGCGGAGTACCGGCTTTCGCTCC	551
QY	3055	CTTCCACATCCATGAAGCGGAGTACCGGCTTTCGCTCC	3093
DB	552	CTTCCACATCCATGAAGCGGAGTACCGGCTTTCGCTCC	590

[illegible]



QY 73 GAUCCATCAGATTTCGGATGCCAATATACAAACAGATCCGGTGTGTGTGGGCAAC 132  
DB 112 CCAGCGCTTAGCGTGGCGCCAGAGACTATGAGACATATGCGGTGTTGTGGAGCC 172  
QY 133 AAGCAGAGAGGAGACAGC-----ACAGAGACAGAGCTGGAGATCAGATATATG 184  
DB 172 GGGCGGAGCGCTGAGCGCGGAGAGAGTGTGAGAGATCAAGATCAGAGAGATGCTG 231  
QY 187 ATGATGAGGAGAGCGCTGATGATGTGTGTAGGAGATATTACATGTTGATATAGAC 245  
DB 232 CCGGTACAGAGAGAGCTGTTCATTCGGAGAGAGAGAGATTCAGGAGATGAGATGAG 291  
QY 247 ACATACACACAGGAGAGAAATTATTTAGCAAAAGACTGACATGGAATATGAGAGGCG 306  
DB 292 CCAGCGAGCTCCAGCGAGGTGGGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 351  
QY 307 GATGTAGACACATGCAG 366  
DB 352 GACATAAGCTGTGTGGATGAGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 411  
QY 367 GTTCTTTAAAG 426  
DB 412 GTGCTGCTCTCGGAG 471  
QY 427 TCGTGGAGAGAGATAG 486  
DB 472 GTGTGGCGCAACTACAGCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 531  
QY 487 GCGAGATGCGCATATGATCCAAACATGCGCAAGTTCGCACTGTTTGGAGATGAGAGAG 546  
DB 532 GCGCGCTGCGCTAG 591  
QY 547 TACTCAGCAG 606  
DB 592 TTCACAGCTACTGTACCGGTCTCCTAGGAGATGATGCTGTGATACCGAGAGAGAGAG 651  
QY 607 GAAAGCCCTACCGTGGGAGCGTCAAGCAGAGATCAAAATGATGAGAGAGAGAGAGAG 666  
DB 652 GACAGGCGCACCTGGGCACTGGAACATGATCCAGAGAGAGAGAGAGAGAGAGAGAG 711  
QY 657 GTTCAGCGCTGGATACGAGAGATATATCTACTTCTTGTAGAGAGAGAGAGAGAGAG 726  
DB 712 GTCCATCGGTGGAG 771  
QY 727 TATACAGCATGGAAGAGTATGTTTCCCAAGAGTGTCCAGGTGTGAGAGAGAGAGAG 786  
DB 772 TTATACACTGGAG 831  
QY 787 GGAGGATCTCAAGAGTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 846  
DB 832 GGAGGCTCCCGCGGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 891  
QY 847 TGTGACTGCTGGAG 906  
DB 892 TGGCTGTACCGGAG 951  
QY 907 ATTGCTATCAAG 966  
DB 952 GTCAGCGTGGGCGCGCGCGGTGTGCGCGCTTTTTCAGAGAGAGAGAGAGAGAGAG 1011  
QY 967 CCGTGGGTTCAG 1026  
DB 1012 CCGTGGGTTCAG 1071  
QY 1027 TTCAGGAG 1086  
DB 1072 TTCAGGAG 1131  
QY 1087 AAGCGAG 1146  
DB 1132 CGAGCGCGCGCGCGGTGCT-----CGGAGCGCGCGGAGAGAGAGAGAGAGAGAG 1185  
QY 1147 GAGTTCCTGATGATACCTTGAATCTCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1206

DB 1186 GCGTGGCGAGTACATCTCAACTTTGTCAAGACCCACCTCTGATGACAGCGGCTG 1245  
QY 1267 CCGTCCATCTTCAACAG 1266  
DB 1246 CCGTCCGCTGGCGCATCGCGCTGGATGCTTGGAGAGAGAGAGAGAGAGAGAGAGAG 1305  
QY 1267 ATTGAGTGGAG 1326  
DB 1306 GTGGCTGAGCTGGAG 1365  
QY 1327 GAG 1380  
DB 1366 GAGCGGGAG 1425  
QY 1381 AATGAG 1434  
DB 1425 GCGCTGAGTGTCTTGGAG 1485  
QY 1435 TATGATGAG 1494  
DB 1486 GCGGAG 1545  
QY 1495 CTGTATGCTTGGCTTCTGAG 1554  
DB 1546 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1605  
QY 1555 GGAAGTGTAAAG 1614  
DB 1506 TCGGAGTGTATGAAG 1665  
QY 1515 GGTGCTGCTTGGAG 1674  
DB 1566 GGCT---CGTGCATCTTCCACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1722  
QY 1675 CGTGGCAATACAGATGCTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1703  
DB 1723 GGGCCAG 1751

## RESULT 14

US-10-035-977-34

Sequence 34, Application US/10035977

Publication No. US2003013427A1

## GENERAL INFORMATION:

APPLICANT: Desnoyers, Inc.

APPLICANT: Eaton, Dan L.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Guiney, Austin L.

APPLICANT: Pan, James

APPLICANT: Stewart, Timothy A.

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William L.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3030R1C10

CURRENT APPLICATION NUMBER: US/035,977

CURRENT FILING DATE: 2001-12-26

PRIOR APPLICATION NUMBER: 60/085579

PRIOR FILING DATE: 1998-05-15

PRIOR APPLICATION NUMBER: 60/112514

PRIOR FILING DATE: 1998-12-15

PRIOR APPLICATION NUMBER: 60/113300

PRIOR FILING DATE: 1998-12-22

PRIOR APPLICATION NUMBER: 60/113430

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: 60/113605

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: 60/113621

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: 60/114140

[illegible]



```
DB 652 GACAGGCGCCACCCCTGCGGACCCGCTGAAACATGACTCCAAAGTGGTTCRAAGAGCCCTTACTTT 711
QY 667 GTTCAAGCGGTGGATTACGGAGATTATATCTACTCTCTTCAAGGAAATAGKAGTGGAG 726
DB 712 GTCCAATGGGTGGAGTGGGACGAGCATGCTACTCTCTTCCGGAGACTGGCATGGAG 731
QY 727 TATAACACCATGSSAAGAGTACTTTCCCAAGAGAGTGTGATGGTTTGAAGATGATAIG 786
DB 772 TTTAACTACCTGGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 831
QY 787 GGAGATCTCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 846
DB 832 GGAGCTCCCGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 891
QY 847 TGCAGTCTGAGAGTCTCTATTTTATTTCAACACTTTCAGAGGAGTTTACASATGIG 908
DB 892 TGTCTGTACCGGAGACTGCTCTATTTTATTTCAACACTTTCAGAGGAGTTTACASATGIG 951
QY 907 ATTCATATCAAGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 966
DB 952 GTACGCTCTGCGGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1011
QY 967 COTGGTCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1026
DB 1012 COTGGTCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1071
QY 1027 TTCAGAGAGAGAGTCTCTGTATTTCCAACTTTCAGAGGAGTTTACASATGIG 1086
DB 1072 TTTCCAGAGAGAGTCTCTGTATTTCCAACTTTCAGAGGAGTTTACASATGIG 1131
QY 1087 AAGCCCGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1146
DB 1132 GACGCGCGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1185
QY 1147 GAGTTCCTGTATGATACCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1206
DB 1186 GCTTGTCCGGAATGACATCTCTCAACTTTCAGAGGAGTTTACASATGIG 1245
QY 1207 CCTCTCATCTTCAAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1266
DB 1246 CCTCTCATCTTCAAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1305
QY 1267 ATTGAGTGGACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1326
DB 1306 GTGGCTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1365
QY 1327 GAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
DB 1366 GAGGCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1425
QY 1381 AATGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1434
DB 1426 GAGGCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1485
QY 1435 TATGAGTGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1494
DB 1486 GAGGCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1545
QY 1495 GTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1554
DB 1546 GTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1605
QY 1555 GAGGCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1614
DB 1606 TCGGCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1665
QY 1615 GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1674
DB 1666 GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1722
QY 1675 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1763
DB 1723 GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1751
```

## RESULT 15

US-10-137-870-543  
Sequence 543, Application US/10137870

Publication No. US20030138883A1

## GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Matthew  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Elivaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gursey, Austin D.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Wataabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C155  
CURRENT APPLICATION NUMBER: US/10/137,870  
CURRENT FILING DATE: 2002-05-03  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 543  
LENGTH: 3721  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-137-870-543

## Query Match 17.9% Score 553.8; DB 12; Length 3721;

Best Local Similarity 60.0%; Pred. No. 2.1e-162;  
Matches 1025; Conservative 0; Mismatches 657; Indels 27; Gaps 5;

```
QY 13 GCCTTGTGTGTATTTACACACTGTACACTTTTCTGGGCTGTTCCTCCAGAGATTCT 72
DB 52 GCCTCTCTGCTTCTGTGTGTGTACTGGGGGGCCGACGGGCTCTTCTGAGGAGCG 111
QY 73 GAGCCATCATGTATTTTGTGATGGCACTATACAAACAGTATCGGTGTTTGTGGCCAC 132
DB 112 CCSCCGCTTAGCGTGGCCCGGAGGACTACCTGAACCACTATCCCGTGTGTGGGAGC 171
QY 133 AAGCCAGGAGGAACACAC-----ACAGAGGACAGGGCTGGACATCCAGATTATG 186
DB 172 GGGCCCGGAGCCGACCGCCGAGAGGTGCTGACGACCAACATCCAGAGTCCG 231
QY 187 ATCATGAAGTACAGCTCTACATTGTCTGAGGACCATATTTATCTGTTCATAGAC 246
DB 232 CGGTCAACAGGAGCGTTCATTGGGACAGGAGCAACCTCTACCGGTAGAGCTGGAG 291
QY 247 ACATCACACAGGAGAGAAATTTATGTAGCAAAATGACATGGAATCTACAGAGGC 306
DB 292 CCCCCCGCTCCAGGAGCTGGGTACCGAGGAGCTGACCTGGAGATCTAACCCGAGC 351
QY 307 GATGTACACATGACGAATGAAGGAAAAACATAAGGATGAGTGGCAACATTTTATAA 366
DB 352 GACATAACGTGTCTCGATGAGGGCAACAGAGGGCGAGTGTCCAACTTCGTAAAG 411
QY 367 GTTCTTTAAAGAAAAAGATGATGATGTTGTTCTGTGGAACATATGCTTCAACCT 426
DB 412 GTGTGTCTCTTCGGGAGAGTCCACCTCTTGTGTGGGTTCACAGCGCTTCAACCG 471
QY 427 TCTGTGAGAACTATAGATGAGTATGATGAGACCACTCGGGGATGAATTCACGGAGT 486
DB 472 GTGTGCCCAACTACAGATAGACACCTCGAGCCCTCGGAGACACACATCAGCGGTATG 531
```





112 CCCCCAGCTCCACGGAGCTCGGTACACAGGAGAGCTGACCTGGAGATCTAACCCAGC 371  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

1381 AATGACAGCCCTTTCCIGGAGGAGATGAGTGTGTACAACTCTGAAATATCCAG-----C 1434  
1446 GGGCICAGTGTCTTCCCTGGAGGAGTGTGAGACCTACCCGCCGACAGGTGTGTGACGGGCC 1505  
1435 TATGATGGAGTCCGAGACAAAGGATCATGGGCATGACAGTGGAGCTGGACAGCAAGCAGTCT 1494  
1506 GCGGTGGGACGACAGGCGGCTGTGAGCTTGGAGCTGGACGAGTCTCGGGGGC 1565  
1495 CTGATGTTGGCTCTCTCTACCTGTGTGATTAAGTGTTCCTTGGCCGGTGTGACAGCAT 1554  
1566 CTCTGCTGCTCTTCCGCGGCTCGCTGTCTCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1625  
1555 GGAAGTGTAAAGAACTGTATGCTCCAGACCCCATATTGTGGATGGATAAAGAA 1614  
1626 TCGGTGTATGAGAGCTGTATGCGCAGTCAGGACCTAGTCGGGTGGGCCCCGAC 1685  
1615 GTGTGCTGCTGAGCATTTATACCCAAACAGCAGTGTGCTTTTGTGACGACATAGAG 1674  
1686 GGTCTCTCTGATCTCTCAAGCGCGGACAGAGCGGCTTTGAGCAGCAGTGTCT 1742  
1675 GTGTGCAATACAGATGCTGTGCGGAGCTG 1703  
1743 GCGGCCAGCAGCTCAAGCTTAGGGGACTG 1771

## RESULT 2

US-09-077-940A-1  
Sequence 1, Application 95/09077940A  
Patent No. 659544  
GENERAL INFORMATION:  
APPLICANT: KIMURA, TORU et al.  
TITLE OF INVENTION: NOVEL SEMAPHORIN 2 AND GENE ENCODING THE SAME  
FILE REFERENCE: 0020-4426P  
CURRENT APPLICATION NUMBER: US/09/077,940A  
CURRENT FILING DATE: 1998-06-25  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: Patenting version 3.1  
SEQ ID NO 1  
LENGTH: 3692  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: (1)..(18)  
OTHER INFORMATION:  
NAME/KEY: CDS  
LOCATION: (19)..(2682)  
OTHER INFORMATION:  
NAME/KEY: 3'UTR  
LOCATION: (2683)..(3653)  
OTHER INFORMATION:  
NAME/KEY: polyA\_site  
LOCATION: (3654)..(3692)  
OTHER INFORMATION:  
US-09-077-940A-1

Query Match 17.7% Score 546.8; DB 4; Length 3692;  
Best Local Similarity 60.3%; Pred. No. 2.1e-136;  
Matches 1006; Conservative 0; Mismatches 632; Indels 30; Gaps 5;

57 TTTCGAGAGATCTTGAGCCATCATGATTTCGTCATGCAACTATACAAACAGTATCC 116  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

251	QY	TACGTGTGATATAGACACAACACACAGGAGAGAAATTTATGTAGAGAAATAACTGACATC	293
279	Db	CCAAAGTAGAATGAGAGCCATCCACAACACAGAGAGTGGGTATACAGGGAGATACCTG	319
291	QY	GAATCTNACAGGCGGATGTACACACATGCNNAAGAGGGAAACATAGAGTAAGTAC	350
339	Db	GGCTCCAAACCAGTGAATCATGTGTGGATGAAGGGAAGCAAGAGGGAGAGTG	369
351	QY	CCACAACTTATTAAACTTCTTAAGNANAACGATATGCCATGTGTTCTGTGTTGGAAC	410
399	Db	TCGGAAGCTTGTCAAGTGTCTCTGCTTCGTGAGNATCCACGCTCTCTGTTGGGCTC	459
411	QY	TAATGCTTCAAGCCTTCCTTCGAAAGAACTATAAATGATACATAGAACATTCGGGA	470
459	Db	CAATGCATTCAAATCCCATGTGTGCCAATTACAGATGGAACNACTGAGCTTCTTGAGA	519
471	QY	TGAATTCAAGCGAATGCCANATGCCATATGATGCCAAACATGGAACGTTCATATGTT	530
519	Db	CAACATCACTGGTATGGGCGCTGGCCCTACGAGCCCAAGCATGCTCCGCTCTT	579
531	QY	TGCAGATGGAAACTATACTCAGCCCAATGTAGTGACITTCCTTGGCATTCACCCAGTAT	590
579	Db	CTCAGATGGGATGCTCTTACAGCCGACGTAAGTGACITTCATAGCATCGACGCTGTAT	639
591	QY	TTACCGAGTCTTTGGAGAAAGCCCTACCTTCGGGACCGCTCAAGCAGGATTCAAAATGTT	650
639	Db	CTACCGTAGCCTTTGGGCAACGGCCCACTCTGCCACAGTAAAGCATGACATCCAAAGTGT	699
651	QY	GAAGAACCATACITTTCTTAAGCCGCGGATTACGGAGATATATCTACTCTCTCTCAG	710
699	Db	TAAAGAGCCATACITTTGTGATCGGTGGAGTGGGAAGCCAGCTACTTCTTCTTCGG	759
711	QY	GUAAATAGCAGTGGAGTATTAACACCATGGGAAGTAGTTTCCCAAGAGTGGCTCAGGT	770
759	Db	GGAGATCGKCATCGAGTTTAACATATTCGAAAGAGTGGTGGTCTCGGTGIGGCCGCTGT	819
771	QY	TTCTAACAATGATATGGAGAGATCTCAAGAGATCTCGAGAAACAGTGGACGTCTGTTCCT	830
819	Db	ATGCAAGATGATGTGGCGCGTCCCAACGGTCTGGAGAAGCAGTGGACTTCTCTCCT	879
831	QY	GAAGGCGCGCTTGAAGTGCCACGTTCTGAGAGACTCTCATTTTATTTCACAAITCTCCA	890
879	Db	GAAGGCGCGCTCAACGTCTCGGTGGCTGGGAGCTCACAGTTCTACTTCAATGTACGTCA	939
891	QY	GGCAGTTACAGATGTGATTCGATCAAGCGCGTGATGTCTCTGCAAGCTTTCTTCTAC	950
939	Db	GGCTGTGACTGGTGTGTGAGCTTGGCGCCCTGCATGATCTCTCTCTCTCTTCAAC	999
951	QY	ACCTTTAAGACGACGCTCTGGTCTGACGCTCTGAGCTATGACATCTTCACATGTGACAG	1010
999	Db	TCCTAGCAACAGCATCCTGCGTCACTCTCTGTGTTCATTAACAAATGGGATAC	1059
1011	QY	TGTTTTTACTGGAGATTCAAGAACAGAGACTTCTCTCACTTCAGCTTGGACACAGCTTCC	1070
1059	Db	TGTTTTTGAAGGCGGCTTCTGGAGGAGAGAACTACCTGAGTCAAGTGGAGGAGAGTGGC	1119
1071	QY	TGATGAACAGTTCCTTAAGSCCGCAGGCGAGTTCTCTGTGTTGAGTCAATCTCTTAGAAG	1130
1119	Db	TGAGGACCAAGTACCACCGCCCGAGGCGCGCTGTGTGTCGCTGCGGTATGCA	1179
1131	QY	ATATGAACCTCCAAATGAGTTCGCTGATATAGCTTGAAGCTTCAATCAAGACGTAACGCT	1190
1173	Db	GTACACAGGATCCAAATGCCCTTCTGTGACAGATGTCGAAGTTTGTAAAGACCCAGCT	1232
1191	QY	CATGCAAGAGGCACTGCCCTCCATCTTTACAGGCGATGTTCTGTGASACATATGTCAG	1250
1233	Db	GATGAGCAAGAGCGGTGCGTCCGTGGGCCACTGCGCTTGGATGTGAGAACTCTGATACG	1292
1251	QY	ATACCGGCTTACCAAAATCCAGTGGACAGAGCTGTGGGCCATATCAGANTACACAGT	1310
1293	Db	GCACCACTGTACCCGAGTGTCTGTGGATGTGGGTGGAGGCCCATGGGCACTCAGACAT	1352
1311	QY	GTTTTTCTCGGATCAGAGAGGGAATCATCTTGAAGTTTCTTTTGGCCAGATAGG	1364

```

DB      1353 AGCTTCCTTGGCTCTGAGTTGGSCACAGTCTCTCAAATTCCTTGGAAGCCCAATGCCAG 1412
QY      1365 AAATAGTGGTGTTCCTTAATGACAGCTTTTCTCTGGAGGAGATGAGTCTTT-----A 1415
DB      1413 TGCTCAGGACACACAGGGCCAGCATCTTTTGGAGGAGTTTGAGACCTACCGGCCAGA 1472
QY      1416 CAACCTGTAATAATGCCAGCTATGATGGATCGGAAGACAAAAGGATCATGGCATGCAGCT 1475
DB      1473 CAGGTGTGGAGATCCAGCAGHGTGTGAGTGGGACAAAGCACTTCGAGCCTGGAGCT 1532
QY      1476 GCACAGACCAAGACAGCTCTCTTTATGTTGGCTTCCTACTGTGTGTAATAAGGTTCCTCT 1535
DB      1533 AGATGCTGCTCTCAGTGGCTGTGGAGAGCTTCCCGCGCTGTGGTGTCTGTCTCTGT 1592
QY      1536 TGGCGCTGTGTAAGACATGGGAGTGTGTAATAAACCTGTATTCCTCCAGACAGCCATA 1595
DB      1593 TGCCGCTGTGAGCTGTACTCGGGTGATGAAGACATGATGGACCAAGATCCATA 1652
QY      1596 TTGCGATGATAAAGGAAGTGTGTCTCCAGCATTTATCCACCAACAGCAGACTSAC 1655
DB      1653 CTGGGCTGGCGCCGCA---TGCTCTCTGCATCTTCCTCAGACCAAGAACCACTGCCAC 1709
QY      1656 TTTTCAGCAGACATACAGCGTGCANTACAGATGGTCTGGGGACATG 1703
DB      1710 GTTTCAGACATGTGTGCGGGCCAGCACCTCTGGCTTAGTGACTG 1757

RESULT 3
US-05-254-594-5
: Sequence 5, Application US/09254594
: Patent No. 656694
: GENERAL INFORMATION:
: APPLICANT: KIMURA, Toru
: TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN Y
: FILE REFERENCE: 0023-4527P
: CURRENT APPLICATION NUMBER: US/09/254,594
: CURRENT FILING DATE: 1999-05-11
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 5
: LENGTH: 2790
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (...)
: OTHER INFORMATION: Tissue Type: Child Brain
: NAME/KEY: CDS
: LOCATION: (1)..(2790)
: OTHER INFORMATION: Identification Method: E
: NAME/KEY: misc_feature
: LOCATION: (...)
: OTHER INFORMATION: Identification Method: P for resulting peptide

US-05-254-594-5
Query Match 14.34; Score 443.6; DB 4; Length 2790;
Best Local Similarity 58.4; Pred. No. 8.3e-109;
Matches #49; Conservative 0; Mismatches 584; Indels 15; Gaps 4;

QY      164 GCGTGGACATCCAGATGATTATGATGAGGAGCCCTCTACATCTGCTAGGGACC 223
DB      179 GCGTGGACTTTCAGAGATTCCTGACCTTGGAACCGGACCTTGCTAGTGGCTGCCCGGGATC 238
QY      224 ATATTATACGTGTGATATACACATCATCACACGGGAAGAAATTTATTGTAGCAAAAAC 283
DB      239 ACCTTTTCTCTTCGATCTTCACGCCAGAGAGAGGGAGGGGCTCGTCCCAACAAGT 298
QY      284 TGACATGGAATCTAGACAGCGCGATGTAGACATGCAGANTGAAGGAAACATAGG 343
DB      299 ATCTA---ACATGGAGGAAGCAAGATGTGGAAACTGTGCTGTACGGGAAAGCTGACGG 355

```

Query Match	Score	EB 4	Length
14.38	443.6	EB 4	2790

Best Local Similarity 58.68; Pred. No. 8.3e-109;

Matches 849; Conservative 0; Mismatches 584; Indels 15

164 GGGTGGACATCCAGATGATTATGATCATGAAAGGAACCTCTACATTGCTGCT

179 GGCTGGACTTTCAGAGATTCCGTGACCTTGAAACCGGACCTTGCTAGTGGCTGCC

224 A G T T A G C G G C T C A T A T A C C C A C C C A C C A A T T T A T T C T A C C

2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841

239 ACGTTTCTCCTTCGATCTTCAAGCCGAGAGAGGGGAGGGCTGGTGCC

2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526

Y T G A C A T G G A A A T C - A G A C A G G C G A T G T A G A C A C A T G C A G A T G A A G G G A A A A

299 ATCTA--ACATGGAGAGCCCAAGATGTGGAGAACTGTGCTGTACGGGGAAC









```
QY 761 TGCGTCAAGTTTGTAGAAAGATATGGAGGATCTCAAAAGACTCTGGAGAAACACTGGA 420
DB 823 TGCGCGGTGTGTAAAGTGAGATGGGTGGGTGACAGAGGCTTGATGATGCGCACTGA 882
QY 821 GGTGCTGCTGAGAGCGCTTGAAGTGTCTAGTCTCTGGAGACTCTCATTTTATTICA 880
DB 863 CATCTTCTCTAGCTGAGGCTCAAGTCTGGTCTGGTGGGACTCTACCTTCACTTIG 942
QY 881 ACATTCTTAGAGTGTAGAGATGTATATATCAAGGAGTGTATGTTCTGGGAA 940
DB 943 ATGCTTACAGTCTCTTAACCTGAGGCTGTGAGCTGTATGAGGAGCTGTGGCTTTG 1002
QY 941 CGTTTCTACACTTATACAGATCTGCTGGCTGCACTGTCTGCTGCTATGAGATCTG 1000
DB 1003 TCTTCACTACTAGACACATAGATCTGCTGGCTGTGCACTGTGGGCTCTACCTAGATG 1062
QY 1001 ACATTGCCACTGTTTACTGGAGATTCAGAGACAGAGATGCTGTGATCCACTGGA 1060
DB 1063 ACATTGAGCTGGCTTGAAGGCAAGTTCAGAGACAGAGATGCTGTGATGGGCTGGA 1122
QY 1061 CACCACTTCTATGAGAGAGTGTCTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1120
DB 1123 CTCTGTGCTGAGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1182
QY 1181 GCGACCGCTGATGATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1240
DB 1243 CACACCCACTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1299
QY 1241 CAATGCTGATACCGCTTACCAAAATGAGAGTGTGAGACAGCTGCTGGGCTATGCA 1300
DB 1300 TGACTAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1359
QY 1301 ATCACACTGTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1360
DB 1360 ATACTACAGTCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1419
QY 1361 TAGGAATAGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1420
DB 1420 GGGGACA---GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1476
QY 1421 CTGAAAAATGCACTGATGATG---AGTGAAGACAAAGAGATGCTGGGCTGCTGCTG 1474
DB 1477 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1536
QY 1475 TGACAGACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1534
DB 1537 TGACACTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1596
QY 1535 TTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1594
DB 1597 TGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1656
QY 1595 ATGTTGGA:GG 1605
DB 1657 ACTGTGGATGG 1667
```

## RESULT 7

US-08-121-713D-57

Sequence 57, Application US/08121713D

Patent No. 5639856

GENERAL INFORMATION:

APPLICANT: Goodman, Corey S.

APPLICANT: Kolodkin, Alex L.

APPLICANT: Mathes, David

APPLICANT: Bentley, David R.

APPLICANT: O'Connor, Timothy

TITLE OF INVENTION: The Semaphorin Gene Family

```
NUMBER OF SEQUENCES: 160
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 248 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/121,713D
FILING DATE: 13-SEP-1993
CLASSIFICATION: 5.14
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,527
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 2854 base pairs
TYPE: Nucleic acid
STRANDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 451..2640
US-08-121-713D-57
```

Query Match 5.4%; Score 166; DB 1; Length 2854;

Best Local Similarity 53.0%; Pred. No. 2.2e-34;

Matches 532; Conservative 0; Mismatches 415; Indels 57; Gaps 6;

```
QY 289 TGAATATCTAGACAGCGCATGTAGACATGACAGATGAAGGAAACATGAAGATGAG 348
DB 709 TGGCACTCGTCAGTGGCCATCGGAGCTCTGCTACTCAAGGGGAGTCAGAGCAGAC 768
QY 349 TGGCACTCTTAAAGTCTCTTAAAGAAACAGATGATGCTGTTGCTGTGA 408
DB 769 TGGCAGAACTACATCCGAGTCTGGGAAATTTGAGCATGACCGCTACTCTGCGGT 828
QY 409 ACTAATCCCTTCAACCTCTCTCTGAGAAACATATAGAT---GGATACATTGGAACCATTC 465
DB 829 AGCAACCCCTATAAGCCACTATGTGGCACTAGCCCTCAAGGATGAGATTATGTTGA 888
QY 466 GGGATGAAATCAGCGCAATGGCCATATGATGCCAATATGCCAAGCTGCA 525
DB 869 GAGAAAGATATAGGAGAGAGATTTGGCCATTTGAGCCTGACCAACAGCACTGCA 948
QY 526 CTGTTTGAAGTGAAGTATATCTAGCACTGATGACTTCTGCTGCTGCTGCTGCTGCTG 585
DB 949 ATATACATGAGGAGAAATTTATACAGACACATGGCAGACTTCTGGAAGTACCT 1008
QY 586 GTCA:TTACCGGAGTCTTGGAGAAAGCCCTGCGGACCGGACAGCAGATTCAAA 645
DB 1009 CTCATATACCGGCTC-----CTCTAAGAACAGAGATCTGACCTCAA 1053
QY 646 TGAATGAAGAACATATCTTGTCAAGCGGTGATGAGGATATATATCTACTCTCTC 705
DB 1054 CAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1113
QY 706 TTCAGGAAATAGCAGTGGAGTATACACATGCGGAAAGTAGTCTTCCCAAGAGTGGCT 765
DB 1114 TTCCGAGAGACTGCTGTGAGTACATCACTGCGGAAAGCTATCTATTCAGAGATGGCC 1173
```

766 CAGGTTTGAAGATGATGAGAGATGTCAGAGAGTCTGCGAGAACAGTGGACCTG 825  
1174 AGAGTCTGAACATGACAGAGGGGGGCTCATAGGTTGTGACAGA---TGGACTTCT 1230  
826 TTCCTGAAGCGGGCTGAPCTGGTCTAGTCTCCCGAGAGCTCTCAATTTTATTCACAGATT 885  
1231 TTTTGAATGACGCTGCAAGTGTCTGCTGCTGAGATATCCATTTTACTTCAATGAA 1390  
886 CTCGAGGAGTACAGATGATTCGATATCAACGGGCTGATCTGTCTCGGCAAC---- 941  
1291 ATTCACTCAACAGTGAGATCATGAGAGAAATATATGCTCAAGTCGAGAACTCATC 1350  
942 -----GTTTCTACACCTATTAACAGCACTCCCGGCTGAGTCTGTGCTTCACT 993  
1351 TAGGGTGTCTTCACGACACCACTGAACTCTATTTGGTGGCTGTCTGTGCTTCACT 1410  
994 ATGCTTGACATCCGAGTGTCTTTACTGGGAGATTCAGAGCAAGAGTCTCTGATGCTC 1053  
1411 ATGAAGTCAATATCTGAGTCTATTTGATGCTGCTCAATTAAGAGCAGAGGATGAATCA 1470  
1054 ACCTGGACACCACTTCTGATGAAGAGTCTCTAAGCCGAGCGAGTGTCTGCTGGC 1113  
1471 ACTGTTGGCAGTCCCAAGCTTAAGTGGCAGACCAAGCTTGACATGCTGT----- 1526  
1114 TCATCTCTTTAAGAGATATGCAACCTTCCCAATGAGTTCCTGATGATACCTGAACTTC 1173  
1527 -----GAATGACATCTGACAC-----TTCTGTGATGTCTGTCAATTT 1566  
1174 ATCAAGACCCCGCTCATGATGAGAGTCTCTAAGCCGAGCGAGTGTCTGCTGGC 1113  
1567 GTAAAGTCAATACATGATGATGAGAGCGCTGAGAGCACTTTTACTCGGCCAATCTC 1625  
1234 CTGAGAACATGGTCAGATACGCTTACCAAAATTTGAGTGA 1277  
1527 ATTCGATCAGCTTACAGTACAGATTTACAAAATAGTGTGA 1670

## RESULT 8

US-08-835-268-57  
Sequence 57, Application US/08833268  
Patent No. 5807826  
GENERAL INFORMATION:  
APPLICANT: Goodman, Corey S.  
APPLICANT: Kolodkin, Alex L.  
APPLICANT: Mathies, David  
APPLICANT: Bentley, David A.  
APPLICANT: O'Connor, Timothy  
TITLE OF INVENTION: The Semaphorin Gene Family  
NUMBER OF SEQUENCES: 100  
CORRESPONDENCE ADDRESS:  
ADDRESSER: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 Hush Street, Suite 3200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent'n Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/835,268  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,713  
FILING DATE: 13-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: B94-002-1  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
TELEX:  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2854 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 451..2640  
US-08-835-268-57

Query Match 5.4%; Score 166; DB 1; Length 2854;

Best Local Similarity 53.0%; Pred. No. 2,2e-34;  
Matches 532; Conservative 0; Mismatches 415; Indels 57; Gaps 6;

QY	289	TGGAATCTAGACAGCGCGATCTAGACACATGSCAGATGAGGGAACATAGAGATGAG	348
DB	709	TGGCACTGCTCAGGTGCCATCGCGAGCTCTCTACCTCAAGGGGAATCAGAGAGCAG	768
QY	349	TCCCAACTTATTAAAGTTCTTCTAAAGAAAACGATGATGCTTTTGTCTGTGGA	408
DB	769	TGCCAGACTACATCCGAGTCTCGGGAATTCAGGATGACCGCTACTCATCTCGGGT	828
QY	409	ACTAATGCCCTTCAACCCCTCTCGGAGAACTATAGAT---GGATACATGGAACCATTC	465
DB	829	AGGAAGCCCTATAAGCCACTATGTCCGACACTACGCCCTCAAGGATGGAGATTGTGTA	889
QY	466	GGGATGATTCAGCGGAATGCCAGATGCCCATATGATGCAACATGCAAGCTTGCA	525
DB	889	GAGGAAGATAAGCGGAGAGGATTTGCCCTTTGACCTGACCAACAGCACTGCA	948
QY	526	CTGTTTGCAGATGGAACATATACCTCAGCCACGACTGCTCTCTGCCATTGACCA	585
DB	949	ATAACAGTGAGGAGCAATTTGATCAGCAGAGTGGCAGACTTCTCTGGAACGACCT	1008
QY	586	GTCACTTACCGGAGTCTTGSAGAAAGCCCTACCTCGGAGCGTCAAGCAGCATTCAAA	645
DB	1009	CTCATATACCGGCGCT-----CCTAAGAACAGAGAGATCTGACCTCAA	1053
QY	646	TGGTGAAGAGACCAATATTTGTTCAAGCGTGGATTAGGAGATATATCTACTTCTC	705
DB	1054	CAATTAATGCTCTTAACCTTTCTAACACATGAGIACAATGATTTATATCTCTCTC	1113
QY	706	TTCAAGGAAATAGCAGTGGATTAACACCATGGAAGTAGTTTTCAGAGAGTGGCT	765
DB	1114	TTCCGAGAGACTGTCTTGAGTACATCACTCGGAAAGGATATCTATTCAGAGCTGCC	1173
QY	766	CAGGTTTGAAGAAATGATGAGAGATCTCAAGAGTCTCGGAGCAACATGAGACCTCG	825
DB	1174	AGAGTCTGAACATGACAAGGCGCGCTCATCAGGTTGGTGACAGA---TGGACTCT	1230
QY	826	TCCGAGAGCGGCTTGANTGCTCAGTCTGAGAGCTCTCATTTTATTCACCAAT	885
DB	1231	TTTTGAATCACTCTGAACCTTCCGCTCGGAGATTAACATTTTACTTCAATGAA	1290
QY	886	CTCCAGGCACTTACAGATGATTCGTATCAACGGCGGATGTGTCTCGGCAAC----	941
DB	1291	ATTCACTCAACAGTGACATCATCAAGGAAATATGCTGGTCAAGTGGAGAACTCATC	1350
QY	942	-----GTTTCTACACCTTAAACAGCATCCCTGGGTCTGAGCTGTGCTGATGAC	993
DB	1351	TAGGGTGTCTTCACGACACCACTGAACTCTATTTGGTGGCTGTCTGTGCTTCACT	1410
QY	994	ATGCTTGACATGCGCAGTCTTTACTGGGAGATTCAGAGCAAGAGTCTCTCTGATTC	1053
DB	1411	ATGAAGTCAATATCTGAGTCAITTCATGCTCCATTTAAGAGCAGAGAACTGAACTCA	1470
QY	1054	ACCTGGACACCACTCTCTGATGAAGAGTTCCTAAGCCGAGCGGAGTGTCTGTGCTGGC	1113



APPLICANT: Matthes, David  
 APPLICANT: Bentley, David R.  
 APPLICANT: O'Connor, Timothy  
 TITLE OF INVENTION: The Semaphorin Gene Family  
 NUMBER OF SEQUENCES: 160  
 CORRESPONDENCE ADDRESS: 160  
 ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP  
 STREET: 268 Bush Street, Suite 3200  
 CITY: San Francisco  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/833,391  
 FILING DATE:  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/121,713  
 FILING DATE: 13-SEP-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Osman, Richard A.  
 REGISTRATION NUMBER: 35,627  
 REFERENCE/DOCKET NUMBER: B94-002-1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415)343-4341  
 TELEFAX: (415) 343-4342  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 57:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2854 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 451..2640  
 CS-08-833-391-57

Query Match 5.4% Score 166 DB 3 Length 2854

Best Local Similarity 53.0% Pred. No. 2,2e-34

Matches 532; Conservative 0; Mismatches 415; Indels 57; Gaps 6;

QY 289 TGGAAATCTAGACAGCCGCAATGACACATGACAGATGACAGGAAACATAGGATAG 348  
 DB 709 TGGCACTGTGAGTGCCCAATGGGAGTCTGTACCTCAAGGGGAAAGTCAGAGAGAC 768  
 QY 349 TGCACAAATTTTAAAGTCTTCAAGAAAGAGATGATGATTTGTCTGTGA 408  
 DB 769 TGGCAACATACATGGAGTCTGGGAAATGAGAGATGAGCGCTACTGTCGGGT 828  
 QY 409 ACTAATGCTTGAACCTTCTGACAGAAATATAGAT---GGATACATGAGACATTC 468  
 DB 829 ACAGAGCCTATAGGCATATGTGGCACTAGCGCTCAAGATGGAGATATATCTTTA 868  
 QY 466 GGGGATGATTCAGGCAATGCCAGATGCCATATGATGGCAACATGGCAAGTTGTA 525  
 DB 889 GAGAAAGATATGAGGAGAGAGATGTGGCTTTGACCTGACCAACAGACACTGCA 948  
 QY 526 CTGTTTGCAGATGGAAATATATCTAGCAGCAGTACTGACTTCTTGGCATTTACGCA 585  
 DB 949 ATATACAGTCAGGACAAATTTATCTACAGCAACAGTGGGAGATTTCTTGGACAGCC 1008  
 QY 586 GTCAATTACGGAGTCTTGGAGAAAGCCCTACCTTGGGAGCGGTCAACGACGATTCAAA 645  
 DB 1009 CTATATACCGGCGCC-----CTCTAAGAACAGAGAGATCTGACCTCAA 1053  
 QY 646 TGGTTGAAGAACCATACCTACTTGTGTCAAGCGGTGATTACGGAGATATATCTACTTCTC 705

DB 1054 CAATTAAATGCTCCCTAACTTGTCAACACAAATGGAGTACAATGATTTATATCTCTTC 1153  
 QY 706 TCCAGGGAATAGAGCATGAGTATACACCATATACACCATGAGAAAGGTAGTTTCCGAGAGTGCT 765  
 DB 1224 TCCGAGAGAGTGGCTGTTGAGTACATCAACTGCGGAAAGGCTATCTATTCAGAGTGGCC 1173  
 QY 766 CAGGTGTGTAAAGATGATATGGAGGATCTCAAGAGTCCCTGGAGAAACAGTGGACGCTG 825  
 DB 1174 AAGATCTGTAAACATGACAAAGGCGCCCTCATCAGGCTGTGACAGA---TGGACTTC 1230  
 QY 826 TCCCTGAGGCGCGCTTGAAGTCTCAGTTCCTGAGACTCTCAITTTTATTTCAAAAT 885  
 DB 1231 TTTTGAATACAGTCTGAACCTGTCCTGCTGAGATTAATCCATTTTACTTCAATGAA 1290  
 QY 886 CTCAGGCGAGTTACAGATGATTCGTATCAACGCGGCTGATGTGCTCTGGCAAC---- 941  
 DB 1291 ATTCAAGTCAACAGTACATCATTAAGGAAATATGTTGTTCAAGTGGAGAACTCATC 1350  
 QY 942 -----GTTTCTACACCTTATACAGCATCCCTGGCTGCTGAGTCTGCTCTATGAC 993  
 DB 1351 TACGGTGTCTTACGACACACAGTGAATCTATGTTGGTGTCTGCTGTTTGTSCCTCAGT 1410  
 QY 994 ATGCTTGACATTTGCCAGTCTTTTACTGGGAGATTCAGGAAACAGCAAGTCTCCTGATTC 1053  
 DB 1411 ATGAAGTCAATCTCAGTTCATTTGATGTCCTATTAAGAGCAGGAAACATGAACTCA 1470  
 QY 1054 ACTTGACACACAGTCTCTGATGAGAGTTCCTAAGCCGAGCCAGCGTGTGTGCTGGC 1113  
 DB 1471 AACTGTTGGCAGTCCCAAGCTTAAAGTGCAGAAACCAAGCGCTGGCAATGTCT---- 1526  
 QY 1114 TCATCTCTCTTAGAAGATATCAACCTCCCAATGAGTTCCCTGATGATACCTCAACTTC 1173  
 DB 1527 -----GAAAGACAGTCTGACAC-----TTCCTGATGCTGTCTGCTCAATTT 1566  
 QY 1174 ATCAAGACGACCGCTCATGATGAGGAGTGGCTCCATCTTCAACAGGCCATGGTTC 1233  
 DB 1567 GTAAAGTCAATACACTGATGATGAGGCGGTGCGCCAGCAITTTTACTCGGCCCAATGCT 1626  
 QY 1234 CTGAGAACATGTGACAGATACCGCTTACCAAAATTTGAGTGA 1277  
 DB 1627 ATTGGATCAGCTTACAGTACAGANTTACAAAATAGCTGTGA 1670

# RESULT 11

US-09-060-610-57  
 Sequence 57, Application US/09050610  
 Patent No. 6344544  
 GENERAL INFORMATION:  
 APPLICANT: Goodman, Corey S.  
 APPLICANT: Kolodkin, Alex L.  
 APPLICANT: Matthes, David  
 APPLICANT: Bentley, David R.  
 APPLICANT: O'Connor, Timothy  
 TITLE OF INVENTION: The Semaphorin Gene Family  
 NUMBER OF SEQUENCES: 160  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
 STREET: 268 Bush Street, Suite 3200  
 CITY: San Francisco  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/060,610  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/835,268  
 FILING DATE: 09/08/03  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Osman, Richard A.  
 REGISTRATION NUMBER: 36,627  
 REFERENCE/DOCKET NUMBER: 094-002-1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 343-4342  
 TELEFAX: (415) 343-4342  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 57:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2854 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 451..2640  
 US-09-060-610-57

Query Match 5.4% Score 166; DB 5; Length 2854;

Best Local Similarity 53.0%; Pred. No. 2.2e-34;  
 Matches 532; Conservative 0; Mismatches 415; Indels 57; Gaps 6;

Qy 289 TGGAAATCTAGACAGGCGGATGTAGACATATGAGAAATGAGGAAATATAGAGTGA 348  
 Db 709 TGGACATCTGTGAGTGGGATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGG 769  
 Qy 349 TGGACATCTGTGAGTGGGATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGG 408  
 Db 769 TGGACATCTGTGAGTGGGATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGG 828  
 Qy 409 ACTAATGCTTCAAGCTTCTGAGAAATATATATATATATATATATATATATATAT 465  
 Db 829 AGGAAGCTTATAGGACATATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGG 888  
 Qy 466 GGGATGATATACAGGAGTGGGATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGG 525  
 Db 889 GAGAAAGATATAGGAGTGGGATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGG 948  
 Qy 526 CTGTTTGCAGATGGAAT 585  
 Db 949 ATATACATGAGGACAT 1008  
 Qy 586 GTATTTACGAGGATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGG 645  
 Db 1009 CTCATATACGCGGCC-----CTGTAAGAACACAGAGATGAGTGGGATGAGTGG 1063  
 Qy 646 TGGTGAAGAACACAT 705  
 Db 1054 CAATTAATGCTGAT 1113  
 Qy 706 TTCAGGAAATATAGGAGTGGGATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGG 765  
 Db 1114 TTCAGAGAGATGCTGTGAGTGGGATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGG 1173  
 Qy 766 CAGGTTTGTAGATGATATGAGGATGATATATATATATATATATATATATATATAT 825  
 Db 1174 AGATGCTGTAAATATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1230  
 Qy 826 TTCGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 885  
 Db 1231 TTTTGAATCAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1290  
 Qy 886 CTCAGGAGGATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGGGAG 945  
 Db 1291 ATTGAGTGAAGAGTGAATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGGGAG 1350  
 Qy 942 -----GTTTTCACATTTATAGGATGAGTGGGATGAGTGGGATGAGTGGGATGAG 993  
 Db 1351 FACGGTGTCTTACAGACAGGATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGG 1410

Qy 994 ATGTTGACATTTGCCAGTGTGTTTACTGGGAGATTCAGGAAACAGAGATCTCTGATTCC 1053  
 Db 1411 ATGAAGTCAATATCTTGGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1470  
 Qy 1054 ACCUGGACACCATCTCTGATGAGACAGGATGATGAGACAGGATGATGAGACAGGAT 1113  
 Db 1471 AACTGGTGGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1526  
 Qy 1114 TCATCTCTCTTATAGAGATATGCAACCTTCAAGTGTGATGATGATGATGATGATGATGAT 1173  
 Db 1527 -----GAATGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1566  
 Qy 1174 ATCAAGAGGACCGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1233  
 Db 1567 GTAAAGTCAATATCTTGGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1626  
 Qy 1234 CTGACACAAATGGTACAGATACCGCTTACCAAAATTCAGTGGGA 1277  
 Db 1627 ATTCGATCAGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1670

RESULT 12

1CT-US94-10-51A-57  
 : Sequence 57, Application PC/US9410151A  
 : GENERAL INFORMATION:  
 : APPLICANT: The Regents of the University of California  
 : TITLE OF INVENTION: The Somaparin Gene Family  
 : NUMBER OF SEQUENCES: 66  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: FLEHR HOECH TEST ALBERTSON & HERBERT  
 : STREET: 4 Elmwood Center, Suite 3400  
 : CITY: San Francisco  
 : STATE: CA  
 : COUNTRY: USA  
 : ZIP: 94111-4187  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent in Release #1.0, Version #1.25  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: PCT/US94/10151A  
 : FILING DATE: 13-SEP-1994  
 : CLASSIFICATION:  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Osman, Richard A.  
 : REGISTRATION NUMBER: 36,627  
 : REFERENCE/DOCKET NUMBER: FP-58750-PC/RAG  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (415) 781-1989  
 : TELEFAX: (415) 398-3249  
 : TELEX: 910 277299 FHT UR  
 : INFORMATION FOR SEQ ID NO: 57:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 2854 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: double  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: cDNA  
 : FEATURE:  
 : NAME/KEY: CDS  
 : LOCATION: 451..2640  
 : PCT-US94-10151A-57

Query Match 5.4% Score 166; DB 5; Length 2854;

Best Local Similarity 53.0%; Pred. No. 2.2e-34;  
 Matches 532; Conservative 0; Mismatches 415; Indels 57; Gaps 6;

Qy 289 TGGAAATCTAGACAGGCGGATGTAGACATGAGGAAATGAGGAAATGAGGATGAG 348  
 Db 709 TGGACATCTGTGAGTGGGATGAGTGGGATGAGTGGGATGAGTGGGATGAGTGG 769









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Computer Ltd.

CM nucleic - nucleic search, using sw model

Run on: September 30, 2003, 16:34:33 ; Search time 1:02:2.7 Seconds  
(without alignments)  
11427.122 Million cell updates/sec

Title: US-09-856-681-1

```

Perfect score: 3093
Sequence: 1 atgaatcagaagccctct.....ccatctatgaatctaca:aa 3094

```

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext. 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post-processing: Minimum Match: 0%  
Maximum Match: 100%  
Listing first 45 summaries

Database : GenEmbl;♦

1:	qb.ba:
2:	qb.htg:
3:	qb.b:
4:	qb.cm:
5:	qb.ov:
6:	qb.pat:
7:	qb.p:
8:	qb.pl:
9:	qb.pr:
10:	qb.ro:
11:	qb.sts:
12:	qb.sy:
13:	qb.un:
14:	qb.vt:
15:	em.ba:
16:	em.h:
17:	em.hum:
18:	em.in:
19:	em.mu:
20:	em.cm:
21:	em.or:
22:	em.ov:
23:	em.pat:
24:	em.ph:
25:	em.pl:
26:	em.ro:
27:	em.sts:
28:	em.un:
29:	em.vt:
30:	em.htg.hum:
31:	em.htg.in:
32:	em.htg.other:
33:	em.htg.mu:
34:	em.htg.pl:
35:	em.htg.ro:
36:	em.htg.mam:
37:	em.htg.vrt:
38:	em.sy:
39:	em.htg.hum:
40:	em.htg.mu:
41:	em.htg.other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Ro.	Score	Query			ID	Description
			Match	Length	Pos		
C	1	3093	100.0	3093	6	AX026741	Sequence
	2	3093	100.0	3862	6	AX026746	Sequence
	3	3093	100.0	3862	9	AF279656	Homo sapi
	4	3032	98.0	4250	9	AB037789	Homo sapi
	5	2969.6	96.0	6065	6	BC160721	Primer fo
	6	2969.6	96.0	6060	9	BC027867	Homo sapi
	7	2746.2	88.8	3550	6	AX096520	Sequence
	8	2414.8	78.1	3018	10	AF288866	Mus muscu
	9	2182.4	70.5	3041	9	BC156617	Primer fo
	10	2182.4	70.5	3041	9	BC027501	Homo sapi
C	11	2137.4	65.1	2770	10	AF030430	Mus muscu
	12	1824.4	59.0	2293	6	ED157826	Primer fo
	13	1824.4	59.0	2293	9	AK027471	Homo sapi
	14	1742.4	56.3	2227	6	BC127304	Primer fo
	15	1742.4	56.3	2227	9	AK074975	Homo sapi
	16	1666.4	53.9	2306	6	BC159853	Primer fo
	17	1662.4	53.9	2306	9	AK027654	Homo sapi
	18	1462	47.3	2123	6	BC155987	Primer fo
	19	1462	47.3	2123	9	AK074743	Homo sapi
	20	1367.2	44.2	1634	9	AK096337	Homo sapi
C	21	1260	38.6	11823	9	AC010457	Homo sapi
	22	1260	38.6	134349	9	AC010296	Homo sapi
	23	1200	38.6	154061	9	AC008524	Homo sapi
	24	1200	38.6	179647	2	AC0108124	Homo sapi
	25	1200	38.6	188207	2	AC010233	Homo sapi
	26	1023.8	33.1	170366	2	AC124181	Mus muscu
	27	1023.8	33.1	184366	10	AC121783	Mus muscu
	28	1023.8	33.1	237974	2	AC094771	Rattus nc
	29	1023.8	33.1	253924	2	AC094771	Rattus nc
	30	1023.8	33.1	256785	2	AC125773	Rattus nc
C	31	758.6	24.5	839	6	BC147394	Primer fo
	32	758.6	24.5	814	6	BC125056	Primer fo
	33	758.6	24.5	814	6	BC127012	Primer fo
	34	751.2	24.3	780	6	BC147572	Primer fo
	35	742.6	24.0	788	6	BC126945	Primer fo
	36	681.8	22.0	6161	10	AK122515	Mus muscu
	37	672.2	21.7	5886	9	AF389427	Homo sapi
	38	672.2	21.7	5943	9	AF389428	Homo sapi
	39	672.2	21.7	6111	9	AF389429	Homo sapi
	40	669	21.6	4234	6	AX492975	Sequence
C	41	661.6	21.4	1205	6	AX605967	Sequence
	42	559	21.3	5925	9	AB040912	Homo sapi
	43	659	21.3	5925	9	AF389430	Homo sapi
	44	657.4	21.3	662	6	BC150362	Primer fo
	45	642	21.3	662	6	AX605969	Sequence

## ALIGNMENTS

SEQUENCE ID	AX026741
AUTHOR	DOPES
DEFINITION	Sequence from Patent WO931252.
ACCESSION	AX026741
VERSION	AX026741.1
KEYWORDS	GI:10187886
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	Klostermann A. and Behl C. Human semaphorin 6a-1 (semata-a), a gene involved in neuronal development and regeneration mechanisms during apoptosis, and its

use as a potential drug target  
 Patent: WO 0031252-A 1 02-JUN-2000  
 KLOSTERMANN ANDREAS (DE) : MAX PLANCK ZSELUS:HAUS (DE) : BERL  
 CHRISTIAN (DE)

# FEATURES

Location/Qualifiers

1..3093  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

1..3093  
 /note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC9107.1"

/db\_xref="GI:10167887"

/translation="MRPALLLYTLHFAGAGPEPSEPTISLHONYTQYFVYGH  
 KPRNTQHRLLDIQIMIKNGTLYIAARDIYVTDIDTSHIEIYCSKLLIKWSQA  
 DVTGCRMGKGDCHNEPIKLVLLKNDLDALEVCGINAPNPSCRNYKMDLTPEPDDFS  
 GMARCPYDAKHANVALFADGKLYSATVTDALDAIVYRSLGESPTLRTVKHDSKWLK  
 RPFVQAVDYGDIYVFFREIAVENTMCKVYEPVRAQVCKNDMGSSORVLEKOWSF  
 LKARLNCVPQDSHEVENILQAVTDVIRINGROVVLATFEPYNSIPGSAYCAIMLD  
 IASVTGRFKCKSPDS:WTPVPERVPRPGCCAGSSSLERYATSNBEPDITLRFI  
 KTHPLMDKAVPSIFRPMFLKIMVRYRLTKJAVDTAAGPYQNHVPLGSEKGTILKF  
 LARICNSGFUNDLSLFLREMSYINSEKSYDGVEDKR:MGKGLDRASSLYVAFSTV  
 KYPLECRHKGKCKTASNDPYGWIKEGAGSHLSPNSRLTFEGDIERGNTDZIG  
 DCHNSFVALHSSSELPSITTSOSTAGEVSGRGLMDKMLHLLDPSDIPLGAVSS  
 HHQPKGVIRESLYKHDQVLPVTLALAVLAFVMAVPSGLIYVVCVCHRKQVA  
 VVORKEKELTHSRSGSSSVTKISGLFCTQSKOPKPEALILPLMHNKKILATPQNAK  
 MLIKADQHHLLILATPESIFITQKKRSGSKWERNKLNILNACTKMPMPGSPV  
 TILPLRPSHSHYVVLPIITQOYOHYVDQPKMSVQAOMALEDDAATI EYKTK  
 EHLCSNHNHGV:VENIDSLFQVYDREASLGPPGASLSQGTLSKRLKXHHSSYGV  
 DYKRSYPTNSLIRHOAT:IKPNTNNSNSHLIRNCFGRGDNPPAPQRFVSLQVR  
 SSQPSQCAVTVSRQPSLNAYNSLIRSLKRLTSLKLPDVPVPPKPSFAPLSTSMKPNDACT

RACE COUNT 813 a 855 c 765 g 660 t

## ORIGIN

Query Match 100.0% Score 3093; DR 6; Length 3093;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3093; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCTCAGAGCTTCTGCTATATCCACAGTGGIACACTTTGCGGGGTTGCTTC 60  
 Db 1 ATGAGCTCAGAGCTTCTGCTATATCCACAGTGGIACACTTTGCGGGGTTGCTTC 60

QY 61 CCAGAGATTTCTGAGCCATACACTATCTGCGATGCAATATACAAAACATATTCAGT 120  
 Db 61 CCAGAGATTTCTGAGCCATACACTATCTGCGATGCAATATACAAAACATATTCAGT 120

QY 121 TTGTGGCCACAGCCAGAGGAGACACACACAGAGGACAGATTCAGATTCAGATTC 180  
 Db 121 TTGTGGCCACAGCCAGAGGAGACACACACAGAGGACAGATTCAGATTCAGATTC 180

QY 181 ATTATGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
 Db 181 ATTATGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240

QY 241 ATGACACATCAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300  
 Db 241 ATGACACATCAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300

QY 301 CAGGCGGATGTACACATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 360  
 Db 301 CAGGCGGATGTACACATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 360

QY 361 ATTAAGATTTCTTAAGAAACAGATGATGATGATGATGATGATGATGATGATGAT 420  
 Db 361 ATTAAGATTTCTTAAGAAACAGATGATGATGATGATGATGATGATGATGATGAT 420

QY 421 AACCTTCTCTGACAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 480  
 Db 421 AACCTTCTCTGACAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 480

QY 481 GGAATGCCAGATGCGCATATGATGCCAATATGCCAATATGCCAATATGCCAATATG 540

481 GGAATGCCAGATGCGCATATGATGCCAATATGCCAATATGCCAATATGCCAATATG 540  
 541 AAACATATATCAGGACAGAGTCACTATAC:TCCTTCCGATTTACCGAGTCATTTACCGAGT 600  
 541 AAACATATATCAGGACAGAGTCACTATAC:TCCTTCCGATTTACCGAGTCATTTACCGAGT 600  
 601 CTTCGAGAAAGAGCTTACCTTCGGACCTGCAACAGCAGATTCAAANTGGTTGAAGAAGCA 660  
 601 CTTCGAGAAAGAGCTTACCTTCGGACCTGCAACAGCAGATTCAAANTGGTTGAAGAAGCA 660  
 661 TACTTTGTTCAAGCCGTGATACGGAGATTAATATCTACTTCTTCTTCAAGGAAATAGCA 720  
 661 TACTTTGTTCAAGCCGTGATACGGAGATTAATATCTACTTCTTCTTCAAGGAAATAGCA 720  
 721 GTGGAGATTAACACCAATCGGAAAGAGTACTTTTCCCAAGAGTGGCTCAGGTTTGTAGAAT 780  
 721 GTGGAGATTAACACCAATCGGAAAGAGTACTTTTCCCAAGAGTGGCTCAGGTTTGTAGAAT 780  
 781 GATATGGAGAGATCTCAAGAGAGTCTCTGGAGAAACAGTGGAGCTGGTTCCTGAAGGCGGC 840  
 781 GATATGGAGAGATCTCAAGAGAGTCTCTGGAGAAACAGTGGAGCTGGTTCCTGAAGGCGGC 840  
 841 TTGAACCTGCTCAGTTCCTTGGAGACTCTCATTTTATTTCAACATTC:CCAGGCGAGTTACA 900  
 841 TTGAACCTGCTCAGTTCCTTGGAGACTCTCATTTTATTTCAACATTC:CCAGGCGAGTTACA 900  
 901 GATGATCTGATCAAGCGGCGGTGATGTTGCTGGCAAGCTTCTTACACCTTATTAAC 960  
 901 GATGATCTGATCAAGCGGCGGTGATGTTGCTGGCAAGCTTCTTACACCTTATTAAC 960  
 961 AGCATCCCTGGTCTGCAAGTCTGTGCTCATCATGCTGACATTCGACATTCGACATTCGACATTC 1020  
 961 AGCATCCCTGGTCTGCAAGTCTGTGCTCATCATGCTGACATTCGACATTCGACATTCGACATTC 1020  
 1021 GGGAGATTCAGAGAAACAGAGTCTCTGATTCACCTTCGACACACAGTTCCTGATGAACGA 1080  
 1021 GGGAGATTCAGAGAAACAGAGTCTCTGATTCACCTTCGACACACAGTTCCTGATGAACGA 1080  
 1081 GTCTCTAAGCCGACGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 1081 GTCTCTAAGCCGACGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
 1141 TCCATGATGATTCCTGATGA:ACCTTGAACCTTCATCAAGAGCAGCAGCTCATGATGAG 1200  
 1141 TCCATGATGATTCCTGATGATGA:ACCTTGAACCTTCATCAAGAGCAGCAGCTCATGATGAG 1200  
 1201 GAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 1201 GAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 1261 ACCAAATTCAGTTCGACAGAGTCTGCTGGGCGCATATCAGAAATCACACTGTGCTGCTGCTGCT 1320  
 1261 ACCAAATTCAGTTCGACAGAGTCTGCTGGGCGCATATCAGAAATCACACTGTGCTGCTGCTGCT 1320  
 1321 GATCAGAGAGGAGGATGATCTTGAAGTTCCTTGGCAGAGTGAAGTGAAGTGAAGTGAAGTGA 1380  
 1321 GATCAGAGAGGAGGATGATCTTGAAGTTCCTTGGCAGAGTGAAGTGAAGTGAAGTGAAGTGA 1380  
 1381 AATGACAGCTTTCCTGAGGAGATGATGTTTACAACTTCGAAATGCAGCTATGAT 1440  
 1381 AATGACAGCTTTCCTGAGGAGATGATGTTTACAACTTCGAAATGCAGCTATGAT 1440  
 1441 GGAGTGGAGAGAAAGAGTATCGGCATTCAGCTGGACAGAGCAAGCAGCTCTCTCTAT 1500  
 1441 GGAGTGGAGAGAAAGAGTATCGGCATTCAGCTGGACAGAGCAAGCAGCTCTCTCTAT 1500  
 1501 GTTGCTTCT 1560  
 1501 GTTGCTTCT 1560  
 1561 TGTAAAAAAGCTGTATTCCTCCAGAGAGCAGATTCATTCGATGATGAAGAGAGAGTGGT 1620  
 1561 TGTAAAAAAGCTGTATTCCTCCAGAGAGCAGATTCATTCGATGATGAAGAGAGAGTGGT 1620



Query Match 100.0% Score 3093 DB 6 Length 3862  
Best Local Similarity 100.0% Pred. No. G  
Matches 3093 Conservative C Mismatches G Indels G Gaps G  
BASE COUNT 971 a 1111 c 367 g 813 t  
ORIGIN

1 ATGAGTCAGAGCCCTTGCTGCTAATTTCACACTGCTACACTTGTGCTGGGCTGGTTC 63  
Db 658 ATGAGTCAGAGCCCTTGCTGCTAATTTCACACTGCTACACTTGTGCTGGGCTGGTTC 717  
QY 61 CCAGAAGATCTTACCAATCAATGATTCGATGCAACTATACAAACACATATCCGCTG 120  
Db 718 CCAGAAGATCTTACCAATCAATGATTCGATGCAACTATACAAACACATATCCGCTG 777  
QY 121 TTGTGGCCACAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 280  
Db 778 TTGTGGCCACAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 837  
QY 181 ATATGATCATGAACGGAACCTCTACATTTGTGCTAGGACCATATTATTTACTGTTCAT 246  
Db 838 ATATGATCATGAACGGAACCTCTACATTTGTGCTAGGACCATATTATTTACTGTTCAT 897  
QY 241 ATAGACACATCAACACAGGAGAGATTTATTTACAAACAACTACATGGAATCTAGA 300  
Db 898 ATAGACACATCAACACAGGAGAGATTTATTTATGACAAACAACTACATGGAATCTAGA 957  
QY 301 CAGGCGGATGTAGACACATGATAGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
Db 958 CAGGCGGATGTAGACACATGATAGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1017  
QY 361 ATTAAGTCTCTTCAAGAAACGATGATGATTTGCTGTGGGACATTTGCTTC 420  
Db 1018 ATTAAGTCTCTTCAAGAAACGATGATGATTTGCTGTGGGACATTTGCTTC 1077  
QY 421 AACCTTCTCTGAGAACTATAGATGATGATGATGATGATGATGATGATGATGATGAT 480  
Db 1078 AACCTTCTCTGAGAACTATAGATGATGATGATGATGATGATGATGATGATGATGAT 1137  
QY 481 GGAATGCGAGATGCGCATATGATGCGAAACGATGATGATGATGATGATGATGATGAT 540  
Db 1138 GGAATGCGAGATGCGCATATGATGCGAAACGATGATGATGATGATGATGATGATGAT 1197  
QY 541 AAATATATCAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
Db 1198 AAATATATCAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1257  
QY 601 GTTGAGAAAGGCGTACGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
Db 1258 GTTGAGAAAGGCGTACGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1317  
QY 661 TACTTTGTTCAAGGCTGGATACGAGATATATATCTATCTCTTTCAGGAGGAGGAGGAG 720  
Db 1318 TACTTTGTTCAAGGCTGGATACGAGATATATATCTATCTCTTTCAGGAGGAGGAGGAG 1377  
QY 721 GTGAGTATTAACACCATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
Db 1378 GTGAGTATTAACACCATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1437  
QY 781 GATATGGGAGGATCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
Db 1438 GATATGGGAGGATCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1497  
QY 841 ITCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
Db 1498 ITCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1557  
QY 901 GATGTGATTCATCAACCGGCGGTGATGTGTGCTGCGCAACGCTTTTCTACACCTTAT 960

1558 GATGTGATTCATCAACCGGCGGTGATGTGTCTGCGCAAGCTTTTCTACACCTTATAAC 1617  
QY 961 AGCATCCCTGGCTGTGGAGCTGTGGCTGTATGATGATGATGATGATGATGATGATGAT 1620  
Db 1628 AGCATCCCTGGCTGTGGAGCTGTGGCTGTATGATGATGATGATGATGATGATGATGAT 1677  
QY 1621 GGGAGATTCAGAGAAAGATGCTGCTGATTCACCTGGACACAGTTCCTGATGAACGA 1680  
Db 1678 GGGAGATTCAGAGAAAGATGCTGATTCACCTGGACACAGTTCCTGATGAACGA 1737  
QY 1738 GTTCTAAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1797  
Db 1798 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1857  
QY 1857 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1917  
Db 1917 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1977  
QY 1977 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2037  
Db 2037 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2097  
QY 2097 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2157  
Db 2157 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2217  
QY 2217 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2277  
Db 2277 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2337  
QY 2337 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2397  
Db 2397 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2457  
QY 2457 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2517  
Db 2517 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2577  
QY 2577 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2637  
Db 2637 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2697  
QY 2697 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2757  
Db 2757 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2817  
QY 2817 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2877  
Db 2877 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2937  
QY 2937 TCCATGAGTTCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000

Db 2638 ATGGGCGCGCTCTTCTTCGCGGCAICACCGGTCIACCTGGTCTGTGATCATCGGCGCAAGAC 2697

QY 2041 GTGGCTGTGTGTCAGGCGCAAGAGAGAGTACACCCACTCGCGCGGCGGTCCATGAGC 2100

Db 2698 GTGGCTGTGTGTCAGGCGCAAGAGAGAGTACACCCACTCGCGCGGCGGTCCATGAGC 2757

QY 2102 AGCGTAAAGCTACAGCGGCTCTTTGGGACACACACACACACACACACACACACACACAC 2160

Db 2758 AGCGTAAAGCTACAGCGGCTCTTTGGGACACACACACACACACACACACACACACACAC 2817

QY 2161 GCCATCTCAGCGCACTCAAC 2220

Db 2818 GCCATCTCAGCGCACTCAAC 2877

QY 2221 ATGCTCATTAAGAC 2280

Db 2878 ATGCTCATTAAGAC 2937

QY 2281 ACCCAACGCTGCAGCAGAGCGGAGCGGACCGCGGCGGCGGAGTGGGAGGAGAC 2340

Db 2938 ACCCAACGCTGCAGCAGAGCGGAGCGGACCGCGGCGGCGGAGTGGGAGGAGAC 2997

QY 2341 CAGAACCTCATCAATCCCTGCGCAAAAGACATCGCGGCGGCGGAGTGGGAGGAGAC 2400

Db 2998 CAGAACCTCATCAATCCCTGCGCAAAAGACATCGCGGCGGCGGAGTGGGAGGAGAC 3057

QY 2401 ACGAGCTGCGGCTGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2460

Db 3058 ACGAGCTGCGGCTGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3117

QY 2461 ACGAGCTGCGGCTGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2520

Db 3118 ACGAGCTGCGGCTGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3177

QY 2521 ATGGGCTGAGAGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2580

Db 3178 ATGGGCTGAGAGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3237

QY 2581 AGCAAGAGTCCCAACCATGGGTGAACCTGTGGAGAACTGACAGGCTGCGGCGGCGG 2640

Db 3238 AGCAAGAGTCCCAACCATGGGTGAACCTGTGGAGAACTGACAGGCTGCGGCGGCGG 3297

QY 2641 GTTCCACAGCGGAGCGCTCCCTGGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2700

Db 3298 GTTCCACAGCGGAGCGCTCCCTGGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3357

QY 2701 AGCAAGCGGCTGAACATGACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760

Db 3358 AGCAAGCGGCTGAACATGACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3417

QY 2761 CCCACGAATCCGCTCAGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2820

Db 3418 CCCACGAATCCGCTCAGCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3477

QY 2821 TCCTCCAACT 2880

Db 3478 TCCTCCAACT 3537

QY 2881 CCGCGCGCGGAGAGGTGCACT 2940

Db 3538 CCGCGCGCGGAGAGGTGCACT 3597

QY 2941 GTGACTGTCTCGAGGAGCGCTCAGCGCTCAGCGGCTCAGCGGCTCAGCGGCTCAGCGGCT 3000

Db 3598 GTGACTGTCTCGAGGAGCGCTCAGCGCTCAGCGGCTCAGCGGCTCAGCGGCTCAGCGGCT 3657

QY 3001 AAGGTAAGCGGCTCGGTAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3060

Db 3658 AAGGTAAGCGGCTCGGTAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3717

QY 3061 ACATCATCAAGCGCAATGATGCTGATATAA 3093

Db 3718 ACATCATCAAGCGCAATGATGCTGATATAA 3750

RESULT 3  
AF279656  
LOCUS 3862 bp mRNA linear PRI 11-DEC-2000  
DEFINITION Homo sapiens semaphorin SEMA6A mRNA, complete cds.  
ACCESSION AF279656  
VERSION AF279656.1 GI:11093650  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
SOURCE  
1. 3862  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="3q21-22"  
658..3750  
/note="guidance signal"  
/codon\_start=1  
/product="semaphorin SEMA6A"  
/protein\_id="AA329378.1"  
/db\_xref="GI:11093651"  
/translation="MRSEALLLYFTLLHFGAGGFPEDSEPI:SHGNYTKQYVVFVGH  
KPGNTQRIHRLDIQIMNGTLYIAARDHIYIVDTSHTEITCYCKLTKWKSQA  
DVTCKRMGKHKDECHNFIKVLKNDALFVCGTINAFNPSCRNYKMDLEPFDEBS  
GMACPYDAKHANVALFADGKLSATVTDFAIDAVLYRSLGSEPTLRTVKHDSKVL  
EPYEOADVDCDYIYFEFEIAYENTMGVFRVAOVCCKNDMGSGORVLEKWTSE  
LAKRLNSVPGDSHFYENILQANTDVIRINGROWLATFTIPNSIPGSAYCAYDMLD  
IASVTFRFEKQSPDSTWTFPDERVPRKPPGCCAGSSSLERATISNEEPDDTLNFI  
KTHPMDQAVPSIFNRWELRTWVRYRLTKIADVTAAAGPYONHIVYVFLGSEKGIILKF  
LARIKNGFLNDSLEEMSVYSEKSYDGVEDKIRINGOLDRASSSLYVAFSTCVI  
DCHNSFALNGHSSLLPSTTSDTAQEGVESRGGMLDKWHLDDSPDIDPLGAVSS  
HNHODKGVIRSVLYKGDOLVPVTLAIIVILAVYMGAVESGITYVCVCDHRKQVA  
VYORKEHLTHSRGSSSVTKLSLGEGDTOSKDPKEAILTPLMHNGKLATPGNTAK  
MLTKADQHHLDLALPTPESTPTLOQRKRSRGRERNEONINACTIKMPMPGSPV  
IPTDPLRASPSPHPSVVLPTTQGGYQHEIVDQPKMSVAQMALEDQAALDEYKTK  
FHLSSKPNHGVNLVENDLSIPKVPQREASIGPPGASLSQTLGSLKRLHHSSYGV  
DYKRSYPTNLSLRHQATILKRNNTSSNSHLNRSGFRGNDPPAPQPVDSIQVH  
SSQPSGGQAVIVSRQPSLNAYNSLTRSLKRLTPSLKPDVPPKPSFAPLSTSMKPDNACI  
"

BASE COUNT 971 a 1111 c 967 g 813 t

ORIGIN

Query Match 100.0% Score 3093; DB 9; Length 3862;  
Best Local Similarity 100.0% Pred No. 0;  
Matches 3093; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGTCAGAGCGCTTCCTGCTATATTCACACTGCTACACTTTCGTGGGGTGGTTC 60  
|||||  
658 ATGAGGTCAGAGCGCTTCCTGCTATATTCACACTGCTACACTTTCGTGGGGTGGTTC 7.7  
|||||  
61 CCAGAGATTCGAGCGCAATGATGCTGATATTCGATGGCACTATACAAACAGTATCCGGTG 120  
|||||

Db 718 CCAGAGAGATTCTGACGCGAATCAGTATTTCGATGCGGAAATTAACGAAACAGTATCGGCTG 777  
QY 121 TTTCGGGCGCAACCCAGCGAGCAACACACAGAGAGGACAGGCTCGACATCCGCAAG 130  
Db 778 TTTCGGGCGCAACCCAGCGAGCGGAACACACACAGAGAGGACAGGCTGGACATCCGCAAG 847  
QY 181 ATTATGATCATGAAGAGAGAGGAGGCTCTACATTCGCTGCTAGGCTACCAATATTATATCTGAT 240  
Db 838 ATTATGATCATGAAGAGAGGAGGAGGCTCTACATTCGCTGCTAGGAGGAGGCTCTACATTCGAT 897  
QY 241 ATAGACATATACACACAGGAGGAGGAGGATTTATTTGTAGCAAAAACCTGACATGGAATCTAGA 350  
Db 898 ATAGACATATACACACAGGAGGAGGAGGATTTATTTGTAGCAAAAACCTGACATGGAATCTAGA 957  
QY 301 CAGGCGGATGTAGACATATGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
Db 958 CAGGCGGATGTAGACATATGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1017  
QY 361 ATAAAGTCTTTTAAAGAAACGATGATGCAATGTTTGTCTGGAACCTAATGCGCTTC 420  
Db 1018 ATAAAGTCTTTTAAAGAAACGATGATGCAATGTTTGTCTGGAACCTAATGCGCTTC 1077  
QY 421 ACCCTTCTCGAGAACTATAGATGATGATATGGAACCAATTCGGGATGGAATCAGC 480  
Db 1078 ACCCTTCTCGAGAACTATAGATGATGATATGGAACCAATTCGGGATGGAATCAGC 1137  
QY 481 GAATGGGCGAGTGGCCATATGATGCCAAACATGCCAAGCTTGCACTCTTTGCCAGATGGA 540  
Db 1138 GAATGGGCGAGTGGCCATATGATGCCAAACATGCCAAGCTTGCACTCTTTGCCAGATGGA 1197  
QY 541 AAATATATCTAGCCACACATGACTGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Db 1198 AAATATATCTAGCCACACATGACTGACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1257  
QY 601 CTGGAGAGAGCCCTAGCCTCGGAGCCCTCAGGACCGATTCANAGGCTTGAAGAGCA 660  
Db 1258 CTGGAGAGAGCCCTAGCCTCGGAGCCCTCAGGACCGATTCANAGGCTTGAAGAGCA 1317  
QY 661 TACTTTGTTCAAGGCGGTGATACGAGAGATATATCTACTTCTTCTCAGGAAATAGCA 720  
Db 1318 TACTTTGTTCAAGGCGGTGATACGAGAGATATATCTACTTCTTCTCAGGAAATAGCA 1377  
QY 721 GTGGAGTATAGCCATGGAGAGGTTTCTCCGAGAGGTTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
Db 1378 GTGGAGTATAGCCATGGAGAGGTTTCTCCGAGAGGTTGCTGCTGCTGCTGCTGCTGCTGCTG 1437  
QY 781 GATAAGGAGGATCTCAAGAGGCTGCTGAGAGAGGATGCAATGCTGCTGCTGCTGCTGCTGCTG 840  
Db 1438 GATAAGGAGGATCTCAAGAGGCTGCTGAGAGAGGATGCAATGCTGCTGCTGCTGCTGCTGCTG 1497  
QY 841 TTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
Db 1498 TTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1557  
QY 901 GATGCTG 960  
Db 1558 GATGCTG 1617  
QY 961 AGATCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
Db 1618 AGATCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1677  
QY 1021 GCGAGATTCAGGAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
Db 1678 GCGAGATTCAGGAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1737  
QY 1081 GTTCCTAAGCCAGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
Db 1738 GTTCCTAAGCCAGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1797  
QY 1141 TCCAAATGAGTCCCTGATGATACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
Db 1798 TCCAAATGAGTCCCTGATGATACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1857

QY 1201 GGAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
Db 1858 GGAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1917  
QY 1261 ACCAAATGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1320  
Db 1918 ACCAAATGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1977  
QY 1321 GSATCAGAGAGGAGGATCATCTTCAAGTTTTCGCAAGATAGGAAATAGGAAATAGGAAATAGGAAAT 1380  
Db 1978 GSATCAGAGAGGAGGATCATCTTGAAGTTTTCGCAAGATAGGAAATAGGAAATAGGAAATAGGAAAT 2037  
QY 1381 AATGACAGCCCTTCCGAGGAGGATGAGTGTGTACACTCTGAAATATGAGTGTGTGTGTGTGTGTGT 1440  
Db 2038 AATGACAGCCCTTCCGAGGAGGATGAGTGTGTACACTCTGAAATATGAGTGTGTGTGTGTGTGTGT 2097  
QY 1441 GSAGTGGAGAGCAAAAGATCATGGGCAATGAGGCTGGACAGAGCAAGCAGCTCTCTGTAT 1500  
Db 2098 GSAGTGGAGAGCAAAAGATCATGGGCAATGAGGCTGGACAGAGCAAGCAGCTCTCTGTAT 2157  
QY 1501 GTTGGGTTCTTACCTGT 1560  
Db 2158 GTTGGGTTCTTACCTGT 2217  
QY 1561 TGTAAAAAACCTGTATGCTTCCAGAGACCCATATCTGTGATGATGATGATGATGATGATGATGATGATG 1620  
Db 2218 TGTAAAAAACCTGTATGCTTCCAGAGACCCATATCTGTGATGATGATGATGATGATGATGATGATG 2277  
QY 1621 GCTGAGCAGCTTATACCCAAACAGCAGCTGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680  
Db 2278 GCTGAGCAGCTTATACCCAAACAGCAGCTGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2337  
QY 1681 AATACAGATGGTCTGGGAGCTGCTCAATTCCTTTGGGCACTGAATGGGCACTTCCAGT 1740  
Db 2338 AATACAGATGGTCTGGGAGCTGCTCAATTCCTTTGGGCACTGAATGGGCACTTCCAGT 2397  
QY 1741 TCCCTTCTCCAGCAGCAACACATCAGATTTGGAGGCTCAAGAGGAGGATGATGATGATGATGATGATG 1800  
Db 2398 TCCCTTCTCCAGCAGCAACACATCAGATTTGGAGGCTCAAGAGGAGGATGATGATGATGATGATGATG 2457  
QY 1801 GGAGGAGTCTGGATGGAAGCATCTGCTGACTCACCTGACAGCAGCAGACCCCTTTGGGG 1860  
Db 2458 GGAGGAGTCTGGATGGAAGCATCTGCTGACTCACCTGACAGCAGCAGACCCCTTTGGGG 2517  
QY 1861 GGAGTGTCTCCCAATATCACCAGAGAGAGGAGGATGATTCGGGAAAGTATGATCATAA 1920  
Db 2518 GGAGTGTCTCCCAATATCACCAGAGAGAGGAGGATGATTCGGGAAAGTATGATCATAA 2577  
QY 1921 GGCCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980  
Db 2578 GGCCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2637  
QY 1981 ATGGGGGCGCTTCTTCGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040  
Db 2638 ATGGGGGCGCTTCTTCGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2697  
QY 2641 GTGGCTG 2700  
Db 2698 GTGGCTG 2757  
QY 2701 AGGCTCAGCAAGTCAAGCGGCTCTTTGGGAGCTCAATCAAAAGCCCAAGCCGAG 2760  
Db 2758 AGGCTCAGCAAGTCAAGCGGCTCTTTGGGAGCTCAATCAAAAGCCCAAGCCGAG 2817  
QY 2818 GGCATGCTCAGGCTACTCATGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2877  
Db 2878 GGCATGCTCAGGCTACTCATGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2937

```
QY 2281 ACCCAACGCTCAGCAGAGAGCGGAAGCCAGCGCGGAGCGGCGAGTGGGAGAGGAC 2340
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 2938 ACCCAACGCTCAGCAGAGAGCGGAAGCCAGCGCGGAGCGGAGGAGGAC 2997
QY 2341 CAGAACTCATCAATCCCTGCACAAAGGACATGCCGCCATGGGTGGCTGGATGCC 2400
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 2998 CAGAACTCATCAATCCCTGCACAAAGGACATGCCGCCATGGGTGGCTGGATGCC 3057
QY 2401 ACGGACTGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTG 2460
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 3058 ACGGACTGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTG 3117
QY 2461 ACCGACAGCGGTACGACAGCAGTACGTCGACAGCAGCAGTACGTCGACAGCAG 2520
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 3118 ACCGACAGCGGTACGACAGCAGTACGTCGACAGCAGCAGTACGTCGACAGCAG 3177
QY 2521 ATGGGCTGAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAG 2580
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 3178 ATGGGCTGAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAG 3239
QY 2581 ACCAAGAGTCGCAACCATGAGTACGTCGACAGCAGCAGTACGTCGACAGCAG 2640
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 3238 ACCAAGAGTCGCAACCATGAGTACGTCGACAGCAGCAGTACGTCGACAGCAG 3297
QY 2641 GTTCACAGCGGAGCGCTGCTGGGTGCGCGGAGCTGCTGCTGCTGCTGCTGCTA 2700
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 3298 GTTCACAGCGGAGCGCTGCTGGGTGCGCGGAGCTGCTGCTGCTGCTGCTGCTA 3357
QY 2701 AGCAAGCGCTGCAATGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTA 2760
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 3358 AGCAAGCGCTGCAATGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTA 3417
QY 2761 CCGACAGAGCTGCTGACAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2820
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 3418 CCGACAGAGCTGCTGACAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3477
QY 2821 TCTTCCAAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 3478 TCTTCCAAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3537
QY 2881 CCGCGCCCGCAGAGGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2940
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 3538 CCGCGCCCGCAGAGGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3597
QY 2941 GTGACTGCTGCGAGGAGCGCAGCCTGACAGGCTGACAGGCTGACAGGCTG 3000
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 3598 GTGACTGCTGCGAGGAGCGCAGCCTGACAGGCTGACAGGCTGACAGGCTG 3657
QY 3001 AAGCGTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3060
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 3658 AAGCGTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3717
QY 3661 ACATCCATGAGCCCAATGATGCTGATATAA 3693
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 3718 ACATCCATGAGCCCAATGATGCTGATATAA 3750
```

```
RESULT 4
AB037789
LOCUS      Homo sapiens mRNA for KIAA1368 protein, partial, cds. PRI 14-MAR-2000
DEFINITION
ACCESSION AB037789
VERSION    1
KEYWORDS   Homo sapiens (human)
SOURCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           1 (sites)
REFERENCE  Nagase, T., Kikuno, T., Ishikawa, K.I., Hirose, K.M. and Ohara, O.
           Prediction of the coding sequences of unidentified human genes.
           XVI. The complete sequences of 150 new cDNA clones from brain which
           code for large proteins in vitro
```

```
JOURNAL: DNA Res. 7 (1), 65-73 (2000)
MEDLINE: 20181126
PUBMED: 10718198
REFERENCE: 2 (bases 1 to 4250)
AUTHORS: Ohara, O., Nagase, T. and Kikuno, R.
TITLE: Direct Submission
JOURNAL: Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA technology; 1532-3 Yena, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna@info.kazusa.or.jp,
URL:http://www.kazusa.or.jp/huqe/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
FEATURES             Location/Qualifiers
     source            1..4250
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="f403125"
                     /tissue_type="brain"
                     /cclone_lib="pblucscript11 SK plus"
     gene              1..4250
                     /gene="KIAA1368"
     CDS               <244..3393
                     /note="Start codon is not identified."
                     /codon_start=1
                     /product="KIAA1368 protein"
                     /protein_id="5AA92606.1"
                     /db_xref="GI:7243117"
     ORIGIN
1168 a 1161 c 996 g 925 t
      base count
      OATGTN
Query Match      98.0%; Score 3032; DB 9; Length 4250;
Post Local Similarity 98.4%; Pred. No. G;
Matches 3093; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
QY 1 ATGAGTTCAGAGAGCGCTGCTGCTATATTCACACTGCTACACTTTGCTGGGCGTGGTTTC 60
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 250 ATGAGTTCAGAGAGCGCTGCTGCTATATTCACACTGCTACACTTTGCTGGGCGTGGTTTC 309
QY 61 CCAGAGAGATTCGAGCCCAATCATTTTCGATGGCACTATACAAACAGTATCCGGTG 120
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 310 CCAGAGAGATTCGAGCCCAATCATTTTCGATGGCACTATACAAACAGTATCCGGTG 369
QY 121 TTTGTGGGCCACAGCCAGGAGCGGAACACCACACAGGAGCGACGCTGGACATCCAGTG 180
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 370 TTTGTGGGCCACAGCCAGGAGCGGAACACCACACAGGAGCGACGCTGGACATCCAGTG 429
QY 181 ATTATGATCATGAACGGAACCTCTACATTCCTGCTAGGAGACCATATTATCTGTTGAT 240
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 430 ATTATGATCATGAACGGAACCTCTACATTCCTGCTAGGAGACCATATTATCTGTTGAT 489
QY 241 ATAGACACATACACAGCGAAGAAATTTATTTAGCAAAAACATGACATGAAATCTAGA 300
      |||||..|||||..|||||..|||||..|||||..|||||..|||||..|||||..
Db 490 ATAGACACATACACAGCGAAGAAATTTATTTAGCAAAAACATGACATGAAATCTAGA 549
```







Db 781 TTTCACATCTCCAGACACATACAGATGTGATCGATCAACGGGGTCATCTGTCTT 840  
QY 936 GGCACGTTTCTACACCTTATAACAGATCGCTGGTGTCTGACATCTGTACATGACAT 905  
Db 841 GGCACGTTTCTACACCTTATAACAGATCTGTGGTGTCTGACAGCTGTGCTATGACAT 900  
QY 996 GCTGTACATGTCAGATGTTTATCTGAGATTCAGAGATTCAGAGATGCTGTGTGAC 1055  
Db 901 GCTGTACATGTCAGATGTTTATCTGAGATTCAGAGATTCAGAGATGCTGTGTGAC 960  
QY 1056 CTGGACACCACTCTCTGATGAACAGTTCCTAAGGCCAGGGGAGGTGCTGTGTGCTC 1115  
Db 961 CTGGACACCACTCTCTGATGAACAGTTCCTAAGGCCAGGGGAGGTGCTGTGTGCTC 1020  
QY 1116 ATCCCTCTTAGAAGATATGCAACCTCCATGAGTTCCTCTATGATACCTTGACCTCAT 1175  
Db 1021 ATCCCTCTTAGAAGATATGCAACCTCCATGAGTTCCTCTATGATACCTTGACCTCAT 1080  
QY 1176 CAAGACGACCGGCTCATGGATGAGGAGTTCCTGAGGCTCCCACTTCAACAGGCCATGTTCT 1235  
Db 1081 CAAGACGACCGGCTCATGGATGAGGAGTTCCTGAGGCTCCCACTTCAACAGGCCATGTTCT 1140  
QY 1236 GAGACAAATGGTCAGATACCTCTTACCAAAATGCAATGCAATGCAACAGTTCGTGGGCTATA 1295  
Db 1141 GAGACAAATGGTCAGATACCTCTTACCAAAATGCAATGCAATGCAACAGTTCGTGGGCTATA 1200  
QY 1296 TCAGATACACATGTTGTTTCTGCGATCAGAGAGGGAATCATCTGTGAAGTTTITGGC 1355  
Db 1201 TCAGATACACATGTTGTTTCTGCGATCAGAGAGGGAATCATCTGTGAAGTTTITGGC 1260  
QY 1356 CAGATAGGAATAGTGGCTTCTTAATACAGAGCTTTCTGAGGAGATGAGTGTTA 1415  
Db 1261 CAGATAGGAATAGTGGCTTCTTAATACAGAGCTTTCTGAGGAGATGAGTGTTA 1320  
QY 1416 CAACCTG-AAAAATGCACTATGATGAGTGGAGAGCAAAAGATCATGGCAGCGAGC 1474  
Db 1321 CAACCTGAAATAGTGGCTTCTTAATACAGAGCTTTCTGAGGAGATGAGTGTTA 1380  
QY 1475 TGACAGACAAAGCACTCTTGTATGTGTGTTCTTCTACCTGTGTGATAAGTTTCCC 1534  
Db 1381 TGACAGACAAAGCACTCTTGTATGTGTGTTCTTCTACCTGTGTGATAAGTTTCCC 1440  
QY 1535 TTGGCCGCTGACAGACATGGAAGTGT-AAAAAGCTGTATCTGCTCAGAGACATA 1593  
Db 1441 TTGGCCGCTGACAGACATGGAAGTGTATAAAAAAGCTGTATCTGCTCAGAGACATA 1500  
QY 1594 TATTGTGGATGATTAAGGAGAGTGTGCTGACACCAATTTATCAACCAAGCAGAGATG 1653  
Db 1501 TATTGTGGATGATTAAGGAGAGTGTGCTGACACCAATTTATCAACCAAGCAGAGATG 1560  
QY 1654 ACTTTGACAGACATAGAGTGGCAATAGAGATGCTGAGGAGCTGTCAANTCC 1713  
Db 1561 ACTTTGACAGACATAGAGTGGCAATAGAGATGCTGAGGAGCTGTCAANTCC 1620  
QY 1714 TTGTGGCACTGAATGGGATTCACGTTCCCTCTGGCCAGACCAACCAATCAGATTCG 1773  
Db 1621 TTGTGGCACTGAATGGGATTCGATGCTGGCTGTGAGCAAAACCAATCAGATTCG 1680  
QY 1774 AGGGCTCAAGAGAGATATGAGTCTAGAGAGAGATTCGTAAGAGATCTGTGTAG 1833  
Db 1681 AGGGCTCAAGAGAGATATGAGTCTAGAGAGAGATTCGTAAGAGATCTGTGTAG 1740  
QY 1834 TCACCTGACAGACAGACCTTCTGGGAGAGTGTCTGCGATATACCAACAGAGAGAG 1893  
Db 1741 TCACCTGACAGACAGACCTTCTGGGAGAGTGTCTGCGATATACCAACAGAGAGAG 1800  
QY 1894 GAGTATTCGGGAAAGTTTACCTCAAGGCCACCAACAGCTGCTTCCCGTCAACCTTTG 1953  
Db 1801 GGAGTATTCGGGAAAGTTTACCTCAAGGCCACCAACAGCTGCTTCCCGTCAACCTTTG 1860  
QY 1954 GGCATTGCACTCATCTGCTTCTGATGAGGCGGCTCTCTCGGGGATCAACCGTCTAC 2013  
Db 1861 GGCATTGCACTCATCTGCTTCTGATGAGGCGGCTCTCTCGGGGATCAACCGTCTAC 1920

QY 2014 TCCCTCTGTGATCATCGCGCAAAAGACCTGTGCTGTGCTGAGCGCAAGAGAGATC 2073  
Db 1921 TCCCTCTGTGATCATCGCGCAAAAGACCTGTGCTGTGCTGAGCGCAAGAGAGATC 1980  
QY 2074 ACCCACTTCCGCCGGGGTCTCATGAGCAGCGTCACCAAGCTCAGCGGCTCTTTGGGAC 2133  
Db 1981 ACCCACTTCCGCCGGGGTCTCATGAGCAGCGTCACCAAGCTCAGCGGCTCTTTGGGAC 2040  
QY 2134 ACTCAATCAAAAGCCCAAAAGCCGGAGGCTCTCTCAAGCTGCTCATGCAACAGGCAAG 2193  
Db 2041 ACTCAATCGGAAGCCCAAAAGCCGGAGGCTCTCTCAAGCTGCTCATGCAACAGGCAAG 2100  
QY 2194 CTGGCACTCCGGGCAACACCGGCCAAGTGTCTATTAAAGCAGCAGCAGCAGCTGGAC 2253  
Db 2101 CTGGCACTCCGGGCAACACCGGCCAAGTGTCTATTAAAGCAGCAGCAGCAGCTGGAC 2160  
QY 2254 CTGAGGGGCTCCCGACGCCAGAGTCAACCCCAAGCTGCGAGCAAGCGGAGGCCAGC 2313  
Db 2161 CTGAGGGGCTCCCGACGCCAGAGTCAACCCCAAGCTGCGAGCAAGCGGAGGCCAGC 2220  
QY 2314 CCGCGACCGGGAGTGGAGAGGAGCAGAACCTCATATGCTCTCCAAAGGAGATG 2373  
Db 2221 CCGCGACCGGGAGTGGAGAGGAGCAGAACCTCATATGCTCTCCAAAGGAGATG 2280  
QY 2374 CCGCCCATGGGCTCCCTGTGATTCCCAAGGAGCTGCCCTTGGGGGCTCCCGAGCCAC 2433  
Db 2281 CCGCCCATGGGCTCCCTGTGATTCCCAAGGAGCTGCCCTTGGGGGCTCCCGAGCCAC 2340  
QY 2434 ATCCCGAGGGTGGTGGTCTGCCATCAGCAGAGGGCTACCAAGCATGAGTACGTGGAC 2493  
Db 2341 ATCCCGAGGGTGGTGGTCTGCCATCAGCAGAGGGCTACCAAGCATGAGTACGTGGAC 2400  
QY 2494 CAGCCCAAAATCAGGAGAGTGGCCAGATGGCTGGAGGACAGCGGCCACACTGGAG 2553  
Db 2401 CAGCCCAAAATCAGGAGAGTGGCCAGATGGCTGGAGGACAGCGGCCACACTGGAG 2460  
QY 2554 TATAAGACCTCAAGSAACTCTCAGCAGCAAGAGTCCCAACCATGGGTGAACCTTGTG 2613  
Db 2461 TATAAGACCTCAAGSAACTCTCAGCAGCAAGAGTCCCAACCATGGGTGAACCTTGTG 2520  
QY 2614 GAGACCTTGGAGAGCTGCCCGCCCAAGTCTCCAGCGGGAGGCTTCTTGGTCCCGG 2673  
Db 2521 GAGACCTTGGAGAGCTGCCCGCCCAAGTCTCCAGCGGGAGGCTTCTTGGTCCCGG 2580  
QY 2674 GAGGCTTCTCTCAGACCGGTCTAAGCAAGCGGTGGAATGCAACCTCTCTTTC 2733  
Db 2581 GAGGCTTCTCTCAGACCGGTCTAAGCAAGCGGTGGAATGCAACCTCTCTTTC 2640  
QY 2734 TACGGGTTGACTATAGAGCAGCTTACCCAGCAACTGCTCTCACTCTCCAGAACAG 2793  
Db 2641 TACGGGTTGACTATAGAGCAGCTTACCCAGCAACTGCTCTCACTCTCCAGAACAG 2700  
QY 2794 ACCACTCTCAAAAGSAAAGCAACTAACTGCTCCAAATCTCTCACTCTCCAGAACAG 2853  
Db 2701 ACCACTCTCAAAAGSAAAGCAACTAACTGCTCCAAATCTCTCACTCTCCAGAACAG 2760  
QY 2854 AGCTTTGGCAGGGGAGAGAACCGCGCCCGCCCGCAGAGAGGTGAGCTCCATCAGGTG 2913  
Db 2761 AGCTTTGGCAGGGGAGAGAACCGCGCCCGCCCGCAGAGAGGTGAGCTCCATCCAGGTG 2920  
QY 2914 CACAGCTCCAGGCTATGGCAGGCGCTGACTCTCTGAGCGAGCGCCCAAGGCC 2973  
Db 2821 CACAGCTCCAGGCTATGGCAGGCGCTGACTCTCTGAGCGAGCGCCCAAGGCC 2880  
QY 2974 TACAACCTCACTGACAGGTTCGGGCTGAAGGCTAGCGCTCGCTAAAGCGGAGCTACCC 3033  
Db 2881 TACAACCTCACTGACAGGTTCGGGCTGAAGGCTAGCGCTCGCTAAAGCGGAGCTACCC 2940  
QY 3034 CCAAAAGCATCTTGTCTCCCTTCCACATCCATGAGAGCCCAATGATGGGTGTACATAA 3093  
Db 2941 CCAAAAGCATCTTGTCTCCCTTCCACATCCATGAGAGCCCAATGATGGGTGTACATAA 3000

```

RESULTS 6
AK027867 6360 bp mRNA linear PRI: 01-AUG-2002
LOCUS Homo sapiens cDNA FLJ14961 fis, clone PACE100230, highly similar
DEFINITION tc Mus musculus semaphorin 7a mRNA.
VERSION AK027867.1 GI:14042853
KEYWORDS cDNA capping, fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Iizumi T., Ota T., Hayashi K., Suiyama T., Otsuki T., Suzuki Y.,
Nishikawa T., Nagai K., Sudano S., Takahashi Fujii A., Hara H.,
Tanabe T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
Arita M., Nakamura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
Wakamatsu A., Nakamura Y., Nakafuku K., Masuko Y. and Oshima A.
NEDO human cDNA sequencing project.
TITLE Unpublished
REFERENCE 2 (bases 1 to 6360)
AUTHORS Iizumi T. and Otsuki T.
JOURNAL Direct Submission
TITLE Submitted (10-MAY-2003) Takao Iizumi, Helix Research Institute,
Genomics Laboratory, 1532-3 Yama, Kisarazu, Chiba 252-0812, Japan
COMMENT (E-mail: genomics@hri.co.jp, Tel: 04-38-52-9975, Fax: 04-38-52-3386)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
5'- and 3'-end one pass sequencing and clone selection; Helix
Research Institute (supported by Japan Key Technology Center etc.)
and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
location/Qualifiers
.. 6060
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="P:ACE100230"
/issue_type="placenta"
/clone_lib="vACE4"
/notes="Cloning vector: pMEL3SF13"
89..1345
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAB55418.1"
/db_xref="GI:14042854"
/translating="MIMNGUYIAAROHYIVDID7SHTEIYCSKLLTWKSRQAVD
TCRMKGRKDECNFKVLLKNDALFVCGTNFNPSCRNKMYDLEPPGDFSGMA
RCPYDAKNVALFADGKLYSATVDFLAIDAVIVRSIGESPTLRVKEGSKWKEPY
FYQAVDGYDYTFPRELAVENYTMKVVFVRPOVCKNDGGSORVLEKQWTSFLKA
RNCQSVPGDSHFYFNLCQAVTDV: RINGRDVLA:FTSPYNSIPGSVACVMDLQAS
VLTGRPEKQKSPES:WTPVDERVPKPRCCAGSSSLERATSNRPDDTLNFKTH
PLMDEAVPS:INRPFLE:MYRYRLK:AVDIAAGPYONHIVFVDSKGL:LLFLAR
IGNSQEMDSFLEPMSVINSFKMQL"
BASE COUNT 1639 a 1526 c 1363 g 1532 t
ORIGIN
Query Match 96.0% Score 2469.6 DB 9: Length 6060;
Best local Similarity 99.8% Pred. No. 0;
Matches 2934; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
QY 96 CAACTATACAAACAGTATCCGGTGTCTGTGGCCACAGCCAGCGACGGAACACACACA 155
DB 1 CAACTATACAAACAGTATCCGGTGTCTGTGGCCACAGCCAGCGACGGAACACACACA 60
QY 156 GAGGACACAGGCTGGACATCCAGATGATATCATCATGATGACGACGACCCCTCTACATGCTGCC 215
DB 61 GAGGACACAGGCTGGACATCCAGATGATATCATCATGATGACGACGACCCCTCTACATGCTGCC 120
QY 216 TAGGACACATATTTACTGTTGATATACACATACACACGGAAGAAATTTATGTAG 275
DB 1201 TCAGAAACACACTGTGGTCTTCTGGGATCAGAGAGGGAATCATCTTGAAGTTTGTGGC 1260
DB 121 TAGGACACATATTTATACIGTGTGATATAGACACATACACACGGAAGAAATTTATGTAG 180
QY 276 CAAAAAATGACATGGAATCTTAGACAGCGCGGATGTAGACACATGCGAATGAAGGAAA 335
DB 181 CAAAAAATGACATGGAATCTTAGACAGCGCGGATGTAGACACATGCGAATGAAGGAAA 240
QY 336 ACATAAGGATGAGTCCCAACACATTTTAAAGTCTTCT:AAAGAAACACATGATGCAAT 395
DB 241 ACATAAGGATGAGTCCCAACACATTTTAAAGTCTTCT:AAAGAAACACATGATGCAAT 300
QY 396 GTTGTGTGTGGAACTAATGCCITCAACCCCTTCCTGAGAAACTATAGATGGATACAT 455
DB 301 GTTGTGTGTGGAACTAATGCCITCAACCCCTTCCTGAGAAACTATAGATGGATACAT 360
QY 455 GGAAACCATTCGGGATGAATTCAGCGGAATGGCCAGATGCGCATATGATGCCAAATGC 515
DB 361 GGAAACCATTCGGGATGAATTCAGCGGAATGGCCAGATGCGCATATGATGCCAAATGC 420
QY 516 CAAGCTTGCAGTGTTCGCAGATGGAACACTATCTCAGCCACAGTGCATCTCCTTGC 575
DB 421 CAAGCTTGCAGTGTTCGCAGATGGAACACTATCTCAGCCACAGTGCATCTCCTTGC 480
QY 576 CATTGACGCAATCATTTAGCGGAGTCTTGGAGAAAGCCCTACCCCTGCGGACCGTCAAGCA 635
DB 481 CATTGACGCAATCATTTAGCGGAGTCTTGGAGAAAGCCCTACCCCTGCGGACCGTCAAGCA 540
QY 636 CGATTCAAAATGTTGAAAGAACATACCTTGTTCAGACCGTGGATTCAGGAGATATAT 695
DB 541 CGATTCAAAATGTTGAAAGAACATACCTTGTTCAGACCGTGGATTCAGGAGATATAT 600
QY 696 CTACTTCTCTCAGGAAATACGAGTGGAGTATAATACCATGCGAAAGTAGTTTTCCT 755
DB 601 CTACTTCTCTCAGGAAATACGAGTGGAGTATAATACCATGCGAAAGTAGTTTTCCT 650
QY 756 AAGAGTGGCAGCTTGTAGAAATATATGGAGAGTCTCAAGAGTCTCGGAGAAACA 815
DB 661 AAGAGTGGCAGCTTGTAGAAATATATGGAGAGTCTCAAGAGTCTCGGAGAAACA 720
QY 816 GTGAGCTGCTTCGAGGCGGCTTGAAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 875
DB 721 GTGAGCTGCTTCGAGGCGGCTTGAAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 780
QY 876 TTCAACATCTCTCCAGCAGTGTACAGATGTGATCTGATCAAGGGGGGTGATGTGCTCT 935
DB 781 TTCAACATCTCTCCAGCAGTGTACAGATGTGATCTGATCAAGGGGGGTGATGTGCTCT 840
QY 936 GGCAACCTTTCTACACCTTATACAGCATCCCTGGGTCTGAGTCTGAGTCTGAGTCTGAGT 995
DB 841 GGCAACCTTTCTACACCTTATACAGCATCCCTGGGTCTGAGTCTGAGTCTGAGTCTGAGT 900
QY 996 GCTTGACATTCGCCAGTGTCTTTACTGGGAGATTCAGGAAACAGAGTCTCTGATCTCCAC 1055
DB 901 GCTTGACATTCGCCAGTGTCTTTACTGGGAGATTCAGGAAACAGAGTCTCTGATCTCCAC 960
QY 1055 CTGACACACAGTCTCTGATGAAGAGTCTTAAAGCCAGCCAGGCTGCTGCTGCTGCTGCT 1115
DB 961 CTGACACACAGTCTCTGATGAAGAGTCTTAAAGCCAGCCAGGCTGCTGCTGCTGCTGCT 1020
QY 1116 ATCTCTCTTAGAAGATATGCACTCCATGATGATGATGATGATGATGATGATGATGATGAT 1175
DB 1021 ATCTCTCTTAGAAGATATGCACTCCATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1176 CAAGACACACCGCTCATGATGAGGAGTGGCTCTCATCTTCAACAGGCGCATGGTCTCT 1235
DB 1081 CAAGACACACCGCTCATGATGAGGAGTGGCTCTCATCTTCAACAGGCGCATGGTCTCT 1140
QY 1236 GAGAACATGCTCAGATACCGCTTACCAAAATGCAAGTGGACACAGTCTCTGGGCGATA 1295
DB 1141 GAGAACATGCTCAGATACCGCTTACCAAAATGCAAGTGGACACAGTCTCTGGGCGATA 1200
QY 1296 TCAGAAACACACTGTGGTCTTCTGGGATCAGAGAGGGAATCATCTTGAAGTTTGTGGC 1355
DB 1201 TCAGAAACACACTGTGGTCTTCTGGGATCAGAGAGGGAATCATCTTGAAGTTTGTGGC 1260

```



QY 1 ATGACGTCAGAACCCCTTGCCTGCTATATTCACACTGCTACACTTTCTGGGGCTGCTTC 60  
DB 126 ATGAGCTCAGAACCCCTTGCCTATATTCACACTGCTACACTTTCTGGGGCTGCTTC 185  
QY 62 CCAGAGATTCTGACCCAACTGAGTATTCGCAATGCAACTATACAAAACACTATCCCTG 120  
DB 186 CCAGAGATTCTGACCCAACTGAGTATTCGCAATGCAACTATACAAAACACTATCCCTG 245  
QY 121 TTGTGGGCGACAGCCAGCGGAGGACACACACAGACAGAGCGCTGACATCCAGATG 180  
DB 246 TTGTGGGCGACAGCCAGCGGAGGACACACACAGACAGAGCGCTGACATCCAGATG 305  
QY 181 ATTATGATCATCAAGCGAACCTCTACATTTGCTGTAGGACCATATTTACTCTTGAT 240  
DB 306 ATTATGATCATCAAGCGAACCTCTACATTTGCTGTAGGACCATATTTACTCTTGAT 365  
QY 241 ATAGACATCATCACACGCGAAGAAATTTATTTAGCAAAANAACTGCATGSAATCTABA 300  
DB 366 ATAGACATCATCACACGCGAAGAAATTTATTTAGCAAAANAACTGCATGSAATCTABA 425  
QY 301 CAGGCGGATGTAGACATCCAGATGAGGAGGAAACATAAASATCATGCTGCCACACTT 360  
DB 426 CAGGCGGATGTAGACATCCAGATGAGGAGGAAACATAAASATCATGCTGCCACACTT 485  
QY 362 ATTAAGATTCTCTAAAGAAACAGATCATCATCTTCTGTGTGGCACTTAAGGATCT 420  
DB 486 ATTAAGATTCTCTAAAGAAACAGATCATCATCTTCTGTGTGGCACTTAAGGATCT 545  
QY 421 AGCCTTCTGTGAGAACTATAGATGATGATGATGATGATGATGATGATGATGATGATG 480  
DB 546 AGCCTTCTGTGAGAACTATAGATGATGATGATGATGATGATGATGATGATGATGATG 605  
QY 482 GGAATGGCCAGATGCCATATGATGCCAAGATGCGAAGCTTGGCACTTCTTGCATATGA 540  
DB 606 GGAATGGCCAGATGCCATATGATGCCAAGATGCGAAGCTTGGCACTTCTTGCATATGA 565  
QY 541 AAATATAGTACGACACAGTACTGACTTCCTTGGCATTTGACGAGTCAATTTACCGGAGT 600  
DB 566 AAATATAGTACGACACAGTACTGACTTCCTTGGCATTTGACGAGTCAATTTACCGGAGT 725  
QY 601 CTGGAGAAAGCCCTACCTCGGGACCGTCAAGACAGATTCAAAATGGTGAAGAAAGCA 660  
DB 726 CTGGAGAAAGCCCTACCTCGGGACCGTCAAGACAGATTCAAAATGGTGAAGAAAGCA 785  
QY 661 TACTTTGTTCAAGCGGTGATTAAGGATATATCTACTTCTCTCAGGGAATATGA 720  
DB 786 TACTTTGTTCAAGCGGTGATTAAGGATATATCTACTTCTCTCAGGGAATATGA 845  
QY 721 GTGGAGTATACACCATGGCAAGGTAGTTTCCCAAGAGTGGCTCAGGTTTGTAAAGAT 780  
DB 846 GTGGAGTATACACCATGGCAAGGTAGTTTCCCAAGAGTGGCTCAGGTTTGTAAAGAT 905  
QY 781 GATATGGGAGGATCTCAAGAGCTCTGGAGAACACTGAGCTGCTCTGAGGCGGCG 840  
DB 906 GATATGGGAGGATCTCAAGAGCTCTGGAGAACACTGAGCTGCTCTGAGGCGGCG 965  
QY 841 TTGAATGCTCATCTCTCCAGACTCTCATTTTATTTTCAACATCTCCAGGCACTTACA 900  
DB 966 TTGAATGCTCATCTCTCCAGACTCTCATTTTATTTTCAACATCTCCAGGCACTTACA 1025  
QY 901 GATGTGATTTCGTATACACCGCGGTGAAGTTGTCCTGGCAACGTTTCTACACTTATAAC 960  
DB 1026 GATGTGATTTCGTATACACCGCGGTGAAGTTGTCCTGGCAACGTTTCTACACTTATAAC 1085  
QY 961 AGCATCCCTGGGCTGACAGCTGTGCTATGACATGCTTGCATATGCGAGTCTTTTACT 1020  
DB 1086 AGCATCCCTGGGCTGACAGCTGTGCTATGACATGCTTGCATATGCGAGTCTTTTACT 1145  
QY 1021 GGGAGATTCAAGGACAGAGTCTCTGATTCCACTGACACAGCTTCTCTGATGAAGCA 1080  
DB 1146 GGGAGATTCAAGGACAGAGTCTCTGATTCCACTGACACAGCTTCTCTGATGAAGCA 1205

QY 1381 GTTCTAAGCCCGACGGCCAGCTTGCTGTGCTGGCTCATCTCTCTTAGAACAATATGCAACC 1140  
DB 1206 GTTCTAAGCCCGACGGCCAGGTTGCTGTGCTGGCTCATCTCTCTTAGAACAATATGCAACC 1265  
QY 1141 TCCAAATGAGTTCCTCTGATGATACCTGAATCTCATCAAGACGACACCGCTCATGGATGAG 1200  
DB 1266 TCCAAATGAGTTCCTCTGATGATACCTGAATCTCATCAAGACGACACCGCTCATGGATGAG 1325  
QY 1201 CGAGTGGCTTCATCTCTCAACAGGCCATGGTTCTCTGAGAACAAATGTCAGATACCGCTT 1260  
DB 1326 CGAGTGGCTTCATCTCTCAACAGGCCATGGTTCTCTGAGAACAAATGTCAGATACCGCTT 1385  
QY 1261 ACCAAAATTCAGTGGACACAGCTGCTGGCCATATCAGAAATCACACCTGCTTTTCTG 1320  
DB 1386 ACCAAAATTCAGTGGACACAGCTGCTGGCCATATCAGAAATCACACCTGCTTTTCTG 1445  
QY 1321 GGATCAGAGAGGGAATCATCTTGAAGTTTGTGGCCAGAAATAGGAATATGTTTCTTA 1380  
DB 1446 GGATCAGAGAGGGAATCATCTTGAAGTTTGTGGCCAGAAATAGGAATATGTTTCTTA 1505  
QY 1381 AATGACAGCTTTTCTTGGAGGAGATGAGTGTTCAGAACTCTGAAAATCGACATATGAT 1440  
DB 1506 AATGACAGCTTTTCTTGGAGGAGATGAGTGTTCAGAACTCTGAAAATCGACATATGAT 1565  
QY 1441 GSAATCAACACAAAGAAATCATGGCGATCGACGTCGACAGACAGCAACGACCTCTCTGAT 1500  
DB 1566 GSAATCAACACAAAGAAATCATGGCGATCGACGTCGACAGACAGCAACGACCTCTCTGAT 1625  
QY 1501 GTTCGGTCTCTACTCTGTGTGTAAGGTTCCCTTGGCGGTGTGAACACATGCGGAAG 1560  
DB 1626 GTTCGGTCTCTACTCTGTGTGTAAGGTTCCCTTGGCGGTGTGAACACATGCGGAAG 1685  
QY 1561 TGTAAAANAGCTTATTCCTCCAGAGACCAATTTGTGATGGATGAAGAGAGGTGT 1620  
DB 1686 TGTAAAANAGCTTATTCCTCCAGAGACCAATTTGTGATGGATGAAGAGAGGTGT 1745  
QY 1621 GCCTGCAAGCAATTTATCACCCAAACAGCAGACTGACTTTTGACGAGGACATAGAGCGTGC 1680  
DB 1746 GCCTGCAAGCAATTTATCACCCAAACAGCAGACTGACTTTTGACGAGGACATAGAGCGTGC 1805  
QY 1681 AATACAGATGGTCTGGGAGCTGTCACTANTCTTTGTGGCACTGAATGGSCATTCAGT 1740  
DB 1806 AATACAGATGGTCTGGGAGCTGTCACTANTCTTTGTGGCACTGAATGGSCATTCAGT 1853  
QY 1741 TCCCTCTTGCCACGACACAAACACATCAGATTCGACGGCTCAAGAGGCTATGATCTAGG 1800  
DB 1854 TCCCTCTTGCCACGACACAAACACATCAGATTCGACGGCTCAAGAGGCTATGATCTAGG 1853  
QY 1801 GGAGGAATGTTGAGTGGAGCTGGAGCATCTGCTGACCTCACTTACACAGACAGACCTTGGGG 1860  
DB 1854 GGAGGAATGTTGAGTGGAGCTGGAGCATCTGCTGACCTCACTTACACAGACAGACCTTGGGG 1853  
QY 1861 GCAGTCTTCCCATATATACCAACAGACAGAGGGAGTATTCGGGAAGTATACCTCAAA 1920  
DB 1854 GCAGTCTTCCCATATATACCAACAGACAGAGGGAGTATTCGGGAAGTATACCTCAAA 1880  
QY 1921 GGCACACAGCAGCTGGTTCCTGCTCAGCTCTTGCCCATTTGAGTCTGCTGCTGCTGCTC 1980  
DB 1881 GGCACACAGCAGCTGGTTCCTGCTCAGCTCTTGCCCATTTGAGTCTGCTGCTGCTGCTC 1940  
QY 1981 ATGGGGGCGCTCTTCTCTGGGCAATCAGCTGTACTGCGCTTACTGCTGCTGCTGCTGCTGCT 2040  
DB 1941 ATGGGGGCGCTCTTCTCTGGGCAATCAGCTGTACTGCGCTTACTGCTGCTGCTGCTGCTGCT 2000  
QY 2041 GTGGCTGTGTTGACAGCCAGGAGAGAGCTCACCCACTGGCCCGGGGCTCCATGAGC 2100  
DB 2001 GTGGCTGTGTTGACAGCCAGGAGAGAGCTCACCCACTGGCCCGGGGCTCCATGAGC 2060  
QY 2101 AGGCTCAACAGCTCAAGCGGCTCTTTGGGACACTCAATCAAGACCCAAACCCGAG 2160  
DB 2061 AGGCTCAACAGCTCAAGCGGCTCTTTGGGACACTCAATCAAGACCCAAACCCGAG 2120  
QY 2161 GGCATCTCAGCGCACTCATGACACAGGCAAGCTCCCACTCCCGGCAACACAGCGGCAAG 2220









[illegible]

QY	2890	CAGAGGCTGACATCCAGGTGCACAGCTCCAGCCATCTCGCCAGSSCGTGACTGTC	2940
DB	1981	CAGAGGCTGACATCCAGGTGCACAGCTCCAGCCATCTCGCCAGGCGTGACTGTC	2040
QY	2950	TCAGGCGAGCCGAGCCTCAACGCTCAACCTCACTGCAAGGTGGGGCTGAAGCGTACG	3009
DB	2041	TCAGGCGAGCCGAGCCTCAACGCTCAACCTCACTGCAAGGTGGGGCTGAAGCGTACG	2100
QY	3010	CCCTCGTGAAGCGGGAGCTACCCCGCCAAACCATCTCTCTCCCTTCCTCCATCCATG	3069
DB	2101	CCCTCGTGAAGCGGGAGCTACCCCGCCAAACCATCTCTCTCCCTTCCTCCATCCATG	2160
QY	3070	AGGCCAATGATGGTGATATRA	3093
DB	2161	AGGCCAATGATGGTGATATRA	2184
RESULT 10			
LOCUS	AK027501	3041 bp mRNA linear PRI 01-AUG-2002	
DEFINITION	Homo sapiens cDNA FLJ4595 fis, clone NT2RM4002194, highly similar to Mus musculus semaphorin VIIA mRNA.		
ACCESSION	AK027501		
VERSION	AK027501.1	GI:14042222	
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Naqai,K., Sugano,S., Shiratori,A., Sudo,H., Wagaatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nacahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.		
TITLE	NEDO human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3041)		
AUTHORS	Isogai,T. and Otsuki,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisurazu, Chiba 292-0812, Japan (E-mail:genomics@helix.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)		
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- and 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.		
FEATURES	Location/Qualifiers		
source	1..3041		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="NT2RM4002194"		
	/cell_line="NT2"		
	/cell_type="Teratocarcinoma"		
	/clone_lib="NT2RM4"		
	/note="cloning vector: pUC19FL3-mRNA from uninduced NT2 neuronal precursor cells."		
CDS	85..2184		
	/note="unlabeled protein product"		
	/codon_start=1		
	/protein_id="BAB55158.1"		
	/db_xref="GI:14042223"		
	/translation="MLDIASVFTCRKPKQKSPDSTWTPDSEVPKPRCCACGSSLLRTAISNEFPDSTGLTKIPLMDKAEVPLINRPFWLTWYRLRLTIADVTAGPTQNTVYGLSKDGIILFKELRIGNSGFLINDSLFSEMSYINSEKSYDGVKDKRMGNHLDRASSSLVYFLPKRCRCKGKKTICASRDYCGWIKEGGACSHLSPNRLTFECSIERGNTDGLGDCHNSFVALENGHSSSLIPSTTSDSTAQEGYESRGGMLDWK		

Query Match	70.6%	Score 2182.4	DB #	Length 3042
Best Local Similarity	100.0%	Pred. No. 3		
Matches 219	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY 910	CGTATCAACGGCGGTGATGTTGCTCTGGCAACGTTTCTTACACCTCTATACAGCATCCCT	969		
Db				
Db 1	CGTATCAACGGCGGTGATGTTGCTCTGGCAACGTTTCTTACACCTCTATACAGCATCCCT	60		
QY 970	GGGTCTCAGCTGTGCTGTAIGACATGCTTGACATGCCAGTCTTTTCTACTGAGGATTC	1029		
Db				
Db 61	GGGTCTCAGCTGTGCTGTAIGACATGCTTGACATGCCAGTCTTTTCTACTGAGGATTC	120		
QY 1030	AAGGAACAGAAGTCTCTGTATCCACCTGGACACAGTCTCTGATGAGCAGTCTCTAAG	1049		
Db				
Db 121	AAGGAACAGAAGTCTCTGTATCCACCTGGACACAGTCTCTGATGAGCAGTCTCTAAG	160		
QY 1090	CCGAGGCCAGGTTCCTGTGCTGGCTCATCTCCTTAGAAGATATCGACCTCCATGAG	1149		
Db				
Db 181	CCGAGGCCAGGTTCCTGTGCTGGCTCATCTCCTTAGAAGATATCGACCTCCATGAG	240		
QY 1150	TTCCCTGATGATACCCCTGAATTCATCAAGACGACCGCGCTCATGGATGAGGCACTGCC	1209		
Db				
Db 241	TTCCCTGATGATACCCCTGAATTCATCAAGACGACCGCGCTCATGGATGAGGCACTGCC	350		
QY 1220	TGATATTCACAGAGCCATGTTTCTCTGACAGCAATGTTGATACGCTTCCCAAAAT	1269		
Db				
Db 301	TGATATTCACAGAGCCATGTTTCTCTGACAGCAATGTTGATACGCTTCCCAAAAT	350		
QY 1270	GCAGTGCACACAGCTGGTGGGSCATATCACAATCACACTGTGTTTCTCGGATCAGAG	1329		
Db				
Db 361	GCAGTGCACACAGCTGGTGGGSCATATCACAATCACACTGTGTTTCTCGGATCAGAG	420		
QY 1330	AAGGGAATCATCTGAAATTTTGGCCAGATACGAATPACTGGTTCTTAAATGACAGC	1389		
Db				
Db 421	AAGGGAATCATCTGAAATTTTGGCCAGATACGAATPACTGGTTCTTAAATGACAGC	480		
QY 1390	CTTTCTCTGAGGAGATGAGTGTCTTACACTCTCTGAAAATCGAGCTATGATGAGTGGAA	1449		
Db				
Db 481	CTTTCTCTGAGGAGATGAGTGTCTTACACTCTCTGAAAATCGAGCTATGATGAGTGGAA	540		
QY 1450	GACAAAAGGATCATGGGCATGCGAGCTGGACAGACAGAGCACTCTGTGATGTTGGTTC	1509		
Db				
Db 541	GACAAAAGGATCATGGGCATGCGAGCTGGACAGACAGAGCACTCTGTGATGTTGGTTC	600		
QY 1510	TCTACTGTGTGATAAAGCTTCGCTTGGCGGTGTGACAGCATGGAGTGTAAAAA	1569		
Db				
Db 601	TCTACTGTGTGATAAAGCTTCGCTTGGCGGTGTGACAGCATGGAGTGTAAAAA	660		
QY 1570	ACCTGTATTCGCTCCAGAGACCATATGTTGTAATGATAAAGAGGTGGTGGCTGCAG	1629		
Db				
Db 661	ACCTGTATTCGCTCCAGAGACCATATGTTGTAATGATAAAGAGGTGGTGGCTGCAG	720		
QY 1630	CATTATCACCCACACAGACTACATTTGACAGACAGATAGAGCCGCCAATACAGAT	1689		
Db				
Db 721	CATTATCACCCACACAGACTACATTTGACAGACAGATAGAGCCGCCAATACAGAT	780		
QY 1690	GGCTGGGGGAGCTGCACAACTCTTTTGGCACTCAATGGGATTCACCTTCCTCTTG	1749		
Db				
Db 781	GGCTGGGGGAGCTGCACAACTCTTTTGGCACTCAATGGGATTCACCTTCCTCTTG	840		
QY 1750	CCGAGCACACACCATCATGATTCGAGGGCTTCAGAGAGGATATGAGTCTAGGCGAGGAT	1809		
Db				

DB	841	CCGAGTAAACACCATCATCASAATTCCAGGGCTCAAGAGGGTATATAGTCTAGGGAGGAATG	900
QY	1210	CTGACTGGAAAGCATGTGCTTCACTACCCGACAGCACACACCTTTGGGGCAGTGTCT	1369
DB	901	CTGACTGGAGCATCTGCTTGACTCACTTCACTGACGACACAGACCTTTGGGGCAGTGTCT	960
QY	1370	TCCGTAATCAACAAACAAAGAGGAGTATTGGGAAAGTACTCTCAAGSCCAGGAC	1929
DB	961	TCCGTAATCAACCAAGACAAGAGGAGTATTCGGGAAAGTACTCTCAAGGCCAGGAC	1020
QY	1530	CAGTGGTCCGCTACCTCTCTGGCCATGCGAGTCATCTGGCTTGGCTCAAGGGGCC	1989
DB	1021	CAGTGGTCCGCTACCTCTCTGGCCATGCGAGTCATCTGGCTTGGCTCAAGGGGCC	1080
QY	1990	GTCTTCGGGCATCAACCTCTACTGGGTGTGTATCTCGGGCAAGAGCTGGGTGTG	2049
DB	1081	GTCTTCGGGCATCAACCTCTACTGGGTGTGTATCTCGGGCAAGAGCTGGGTGTG	1140
QY	2050	GTGAGCGAAGGAGAGAGTCAACCACTGTGGCGGGGCTCCATGAGAGCGTCAAC	2109
DB	1141	GTGAGCGAAGGAGAGAGTCAACCACTGTGGCGGGGCTCCATGAGAGCGTCAAC	1200
QY	2110	AAGTCAAGCGGCTCTTTGGGGACATCAATCCAAAGCCCAAGCGGAGGCCATCTCT	2169
DB	1201	AAGTCAAGCGGCTCTTTGGGGACATCAATCCAAAGCCCAAGCGGAGGCCATCTCT	1260
QY	2170	AGCCACTCATGACATGCGAAGCTGGCCACTCCCGGCAACAGGCCCAAGTGCCTCAT	2229
DB	1261	AGCCACTCATGACATGCGAAGCTGGCCACTCCCGGCAACAGGCCCAAGTGCCTCAT	1320
QY	2230	AAAGCAGCAGACACCACTGTGAGCTGACGGGCTCCCAACCCAGAGTCAACCCCAAG	2289
DB	1321	AAAGCAGCAGCAGCAGCTGGACCTCAACGGCTCCCAACCCAGAGTCAACCCCAAG	1380
QY	2290	CTCAGCAGAGCGGAAGCCAGCGCGGCGAGCGCGGAGTGGGAGAGGACCAAGACCTC	2349
DB	1381	CTCAGCAGAGCGGAAGCCAGCGCGGCGAGCGCGGAGTGGGAGAGGACCAAGACCTC	1440
QY	2350	ATCAATGCCCTGCACAAAGGACATGCCGCCATGGGCTCCCTCTGTATTCACAGGAGCTG	2409
DB	1441	ATCAATGCCCTGCACAAAGGACATGCCGCCATGGGCTCCCTCTGTATTCACAGGAGCTG	1500
QY	2410	CGCTGGGGCTCCGTCAGCGACATCCCGAGGGTGTGTCTGTGCCATCAGCGAGCAG	2469
DB	1501	CGCTGGGGCTCCGTCAGCGACATCCCGAGGGTGTGTCTGTGCCATCAGCGAGCAG	1560
QY	2470	GGCTACAGCATGATAGTGTGTGACCGCCCAAAAGACGAGGTGGCCAGATGGCGCTG	2529
DB	1561	GGCTACAGCATGATAGTGTGTGACCGCCCAAAAGACGAGGTGGCCAGATGGCGCTG	1620
QY	2530	GAGCAGCGGTCGCGACATGGAGTATAAGACCATCAAGAGACATCTACGACGAGAGT	2589
DB	1621	GAGCAGCGGTCGCGACATGGAGTATAAGACCATCAAGAGACATCTACGACGAGAGT	1680
QY	2590	CCCAACCATGGGTGAACCTTGTGTGAGAACCTTGACAGCTGCCGCCCAAGTTCACAG	2649
DB	1681	CCCAACCATGGGTGAACCTTGTGTGAGAACCTTGACAGCTGCCGCCCAAGTTCACAG	1740
QY	2650	GAGGAGCGCTCCGTGGTCCCGGGAGCTCCCTGTCTCAGACCGGTCTAAGCAAGCGG	2709
DB	1741	GAGGAGCGCTCCGTGGTCCCGGGAGCTCCCTGTCTCAGACCGGTCTAAGCAAGCGG	1800
QY	2710	CTGGAATGACCACT	2769
DB	1801	CTGGAATGACCACT	1860
QY	2770	TGGCTCACTGAGAGCAACAGCGCCACCTCTCTCAAAAGAAACAACTACTACTCTCAAT	2829
DB	1861	TGGCTCACTGAGAGCAACAGCGCCACCTCTCTCAAAAGAAACAACTACTACTCTCAAT	1920
QY	2830	TGCTCTCACTCTGACAAACAGAGCTTTTGGCAGGGGAGACAACTCCGCGCGCGCGCG	2889
DB	1921	TCCTCTCACTCTGACAAACAGAGCTTTTGGCAGGGGAGACAACTCCGCGCGCGCGCG	1980



DB 1094 GGGAGGTTCAAGGACAGAAATACCCGACATCTTACCTGCAATATACCTGCTTCCAGACGAACGA 1153  
 QY 1091 GTTCTAAGCCAGCCAGGAGTGTGTGCTGGGCTCATCCCTCTAGAAAGATATGGAACC 1149  
 DB 1154 GTGCTAAGCCAGCCAGGAGGCTGTGTGCTGGAATCATCCCTTCTAGAAATATGGAACC 1213  
 QY 1141 TCGAATGATTTCCCTGATATATACCCCTGAGCTTATCATCAAGAGGACACCGGCTCATGGATGAG 1210  
 DB 1214 TCGAATGATTTCCGATGATACCTGATGATTTATTAAGAGGATATGACATCATGAGAG 1273  
 QY 1201 GCAATGCTTCACTTCAACAGAGGATGTTCTGATGAATTAATGTCATAATAGGCTT 1240  
 DB 1274 GCAATGCTTCACTTCAACAGAGGATGTTCTGATGAATTAATGTCATAATAGGCTT 1283  
 QY 1261 ACCAAATTCAGTGGACACAGCTGCTGGGCAATATCAGAAATCACTGTGGTTTTCTG 1320  
 DB 1334 ACCAAATTCAGTGGATAGCAAGCGTCCGGGGCATATCAGAAATCACTGTGGTTTTCTG 1393  
 QY 1321 GGATCAGAGAGGAATCATCTTGAATTTTGGCTAGAAATAGAAATATGCTTTCTTA 1380  
 DB 1394 GGATCAGAGAGGAATCATCTTGAATTTTGGCTAGAAATAGAAATATGCTTTCTTA 1453  
 QY 1381 AATGACAGCTTTTCTGUSAGAGATGAGTGTTTACACACTTAAATGACGCTATGAT 1440  
 DB 1454 AATGACAGCTTTTCTGUSAGAGATGAGTGTTTACACACTTAAATGACGCTATGAT 1513  
 QY 1441 GGATTCAGAGAGGAATCATCTTGAATTTTGGCTAGAAATAGAAATATGCTTTCTTA 1500  
 DB 1514 GGATTCAGAGAGGAATCATCTTGAATTTTGGCTAGAAATAGAAATATGCTTTCTTA 1573  
 QY 1501 GTTCCGTTTCTTACTGTGTATTAAGGTTTCCCTTGGCGGTTGTAACACATATGGGAAG 1560  
 DB 1574 GTTCCGTTTCTTACTGTGTATTAAGGTTTCCCTTGGCGGTTGTAACACATATGGGAAG 1633  
 QY 1561 TGTAAAGAAACCTGTATGCTTCCAGAGNCCATATTGTGATGATATAGAGAGTGGT 1620  
 DB 1634 TGTAAAGAAACCTGTATGCTTCCAGAGNCCGATTTGTGGTGGGTAAAGGAAGTGGT 1693  
 QY 1621 GCTGCGACGCAATTTATCACCATAGAGAGAGTACTTTTGGAGCAGACATAGAGCGTGGC 1650  
 DB 1694 TCGTGTGCCATGTGTACACCCTTATGACAGACTGATGATTTGAGCAGGACATATGACGCTGGC 1753  
 QY 1681 AATACAGATGTCTGGGGATGTGCAATATTCCTTTTGGACAGTAATGGGCACTTCAGT 1740  
 DB 1754 AATACAGAGGCTTAGGAGATGTGCAATATTCCTTTTGGACAGTAATGGGCACTTCAGT 1813  
 QY 1741 TCGCTGTGGCAGCAGCAACACATCATGATTCACGCTTCAAGAGGGGATGAGCTTAGG 1800  
 DB 1814 TCGCTGTGGCAGCAGCAACACATCATGATTCACGCTTCAAGAGGGGATGAGCTTAGG 1873  
 QY 1801 GGAGAAATGTTGGACTGGAAGATCTGATTCATGATGACATGACATGACATGACATGACAT 1840  
 DB 1874 GGAGAAATGTTGGACTGGAAGATCTGATTCATGATGACATGACATGACATGACATGACAT 1933  
 QY 1861 GAGTGTCTTCCCATTAATCAGCAGAGCAAGAGGATGATTCGGGAAACATTAATCTCAAA 1920  
 DB 1934 GAGTGTCTTCTATACAGCAGCAGAGCAGAGAGAGGATGATTCGGGAAACATTAATCTCAAA 1993  
 QY 1921 GGCACAGCAGCAGTGTGTGGGTCACATCTCTTGGCAATTCATTCATCTGAGCTTCTGCT 1980  
 DB 1994 AGCAAGCAGCAGTGTGTGGGTCACATCTCTTGGCAATTCATTCATCTGAGCTTCTGCT 2053  
 QY 1981 ATGGGGGCGCTTCTGAGGATATCAGCTGTCTACTTGGGCTGTGTATCTGAGGAGAAAGAC 2040  
 DB 2054 ATGGGGGCGCTTCTGAGGATATCAGCTGTCTACTTGGGCTGTGTATCTGAGGAGAAAGAC 2113  
 QY 2041 GTGGTGTGGTGCAGCAGAGAGGATGATTCAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 2100  
 DB 2114 GTGGCAGTGTGAGGAGCAGAGAGAGGATGATTCAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 2173  
 QY 2101 AGCTTCAACAGCTCAGCGGCTTTTGGGAGCACTTAATTCAGAGAGCCCAAGAGCTGGAG 2160  
 DB 2174 AGTGTCAACAGCTCAGCGGCTTTTGGGAGCACTTCTTGGGAGCACTTCTTGGGAGCACTTGGAG 2233

QY 2161 GCATCTCTACGCGCACTCATGCAACGCAAGCTGCGCACTCCGCGCAACACAGCGCAAG 2220  
 DB 2234 GCATCTCTACGCGCACTCATGCAACGCAAGCTGCGCACTCCGCGCAACACAGCGCAAG 2293  
 QY 2221 ATGCTCATTAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2280  
 DB 2294 ATGCTCATTAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2353  
 QY 2281 ACCCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2340  
 DB 2354 ACCCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2413  
 QY 2341 CAGAACTCATTAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2400  
 DB 2414 CAGAACTCATTAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2473  
 QY 2401 AGGAGCT 2460  
 DB 2474 AGGAGCT 2533  
 QY 2461 AGGAGCT 2517  
 DB 2534 AGGAGCT 2593  
 QY 2518 CAGATGGCTCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2577  
 DB 2594 CAGATGGCTCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2653  
 QY 2578 ACCAGCAGAGTCCCAACCATGCGGTGAACC 2608  
 DB 2654 AGTAGGAGAGTTCAACCTATGTCTTGAAGC 2684

RESULT 12  
 BD157826  
 LOCUS 2293 bp DNA linear PAT 17-JAN-2003  
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
 ACCESSION BD157826  
 VERSION BD157826.1 GI:27863584  
 KEYWORDS JP 2002191363-A/12669.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 2293)  
 AUTHORS Ota,T., Isoqali,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
 TITLE Primer for synthesizing full-length cDNA and use thereof  
 JOURNAL Patent: JP 2002191363-A 12669 09-JUL-2002;  
 COMMENT HELIX RESEARCH INSTITUTE  
 US Homo sapiens (human)  
 PN JP 2002191363-A/12669  
 PD 09-JUL-2002  
 PE 28-JUL-2009 JP 2002280950  
 PI TOSHIO OTA, TAKAO ISOQALI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORI  
 SAITO,  
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
 PI KEIICHI NAGAI, TETSUJI OTSUKI  
 PC C12N15/09.C07K14/47.C07K16/18.C12N1/15.C12N1/19.C12N1/21.C12N5/ PC  
 10, C12P21/02.C12Q1/48//C12P21/08.G06F17/30.C12N15/00.C12N5/00 CC  
 Primer for synthesizing full-length cDNA and use thereof FH Key  
 Location/Qualifiers  
 FT CDS location(qualifiers  
 1..2293  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 579 a 713 c 581 g 420 t  
 FEATURES  
 source  
 BASE COUNT  
 ORIGIN

Query Match				59.0% Score 1824.4; 23 6; Length 2293;			
Best Local Similarity 99.9%;				Prod. No. 0;			
Matches 1825; Conservative 0;				Mismatches 1; Indels 0; Gaps 0;			
QY	1268	TTGACATGGACAGAGCTGCTGGGCGCATATCAGAAATCACACTGIGTGTTCCTGGGATCAG	1327				
DB	1	TTGACATGGACAGAGCTGCTGGGCGCATATCAGAAATCACACTGIGTGTTCCTGGGATCAG	60				
QY	1328	AGAAGGGAATCACTTCAAACTTTTGGCCACAAATAGTGTCTTCTAAATGACA	1387				
DB	61	AGAAGGGAATCACTTCAAACTTTTGGCCACAAATAGTGTCTTCTAAATGACA	120				
QY	1388	GCCTTTTCCTGGAGGAGATGAGTGTTCAGACTCTGAATAATCAGACTGATGAGACTGG	1447				
DB	121	GCCTTTTCCTGGAGGAGATGAGTGTTCAGACTCTGAATAATCAGACTGATGAGACTGG	180				
QY	1448	AAGACAAAGGAATCATGGGATGACAGCTGGACAGACCAAGCTCTCTGTATGTTCGCT	1507				
DB	181	AAGACAAAGGAATCATGGGATGACAGCTGGACAGACCAAGCTCTCTGTATGTTCGCT	240				
QY	1508	TCCTTACCTGTGTGATAAAAGTTTCGCTTGGCCGCTGTGTGAACACATGCGGAATGAAA	1567				
DB	241	TCCTTACCTGTGTGATAAAAGTTTCGCTTGGCCGCTGTGTGAACACATGCGGAATGAAA	300				
QY	1568	AAACCTGTATGCTCTCAGAGACCCATATCTGTGATGATTAAGCAAGCTGTGGCTGCA	1627				
DB	301	AAACCTGTATGCTCTCAGAGACCCATATCTGTGATGATTAAGCAAGCTGTGGCTGCA	360				
QY	1628	GCCATTATCAGCCACAGAGAGACTGACITTTGAGAGACATAGAGCGGTGGCAATACAG	1687				
DB	361	GCCATTATCAGCCACAGAGAGACTGACITTTGAGAGACATAGAGCGGTGGCAATACAG	420				
QY	1688	ATGTCCTGGGGAGCTGTCACAAATTCCTTTGTCGACTGAAATGGCAATTCAGTTCCTCT	1747				
DB	421	ATGTCCTGGGGAGCTGTCACAAATTCCTTTGTCGACTGAAATGGCAATTCAGTTCCTCT	480				
QY	1748	TGCCACACACACATATCAGAGAGCTGAGAGAGAGATGAGTATGAGTATGAGGAGAA	1807				
DB	481	TGCCACACACACATATCAGAGAGCTGAGAGAGAGATGAGTATGAGTATGAGGAGAA	540				
QY	1808	TGTCGACATGGAGAGATCTCTCTGATGATGATGATGATGATGATGATGATGATGATG	1867				
DB	541	TGTCGACATGGAGAGATCTCTCTGATGATGATGATGATGATGATGATGATGATGATG	600				
QY	1868	CTTCCCAATATCAGCAAGAGAGAGAGAGATGATGAGAGAGATGATGAGAGAGAGAG	1927				
DB	601	CTTCCCAATATCAGCAAGAGAGAGAGAGATGATGAGAGAGATGATGAGAGAGAGAG	660				
QY	1928	ACCAGCTGGTTCGGGTGACATCTCTGAGCAATGACAGTATGCTGGCTTGGTATGGGG	1987				
DB	661	ACCAGCTGGTTCGGGTGACATCTCTGAGCAATGACAGTATGCTGGCTTGGTATGGGG	720				
QY	1988	CGGTCTCTCGGGATCAGAGTCTACGCTGCTGTGATCAGTGGGAGAGAGAGTGGCTG	2047				
DB	721	CGGTCTCTCGGGATCAGAGTCTACGCTGCTGTGATCAGTGGGAGAGAGAGTGGCTG	780				
QY	2048	TGCTGACAGCCAGGAG	2107				
DB	781	TGCTGACAGCCAGGAG	840				
QY	2108	CCAGCTCTACAGCGGCTCTTTGGGAGACATCAATCCAAAGAGAGAGAGAGAGAGAG	2167				
DB	841	CCAGCTCTACAGCGGCTCTTTGGGAGACATCAATCCAAAGAGAGAGAGAGAGAGAG	900				
QY	2168	TCAGGCCATCTATGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2227				
DB	901	TCAGGCCATCTATGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	960				
QY	2228	TTAAGCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	2287				
DB	961	TTAAGCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	1020				

QY	2288	CGCTCAACAGAGCGGAGAGCCAGCCGCGAGCCGCGAGCTGGAGAGAGAGAGAGCC	2347
DB	1021	CGCTCAACAGAGCGGAGAGCCAGCCGCGAGCCGCGAGCTGGAGAGAGAGAGAGCC	1080
QY	2348	TCATCAATGCTGTCACAAAGAGACATGCCCCCATGCGCTCCCTGTGATTTCCACGAGC	2407
DB	1081	TCATCAATGCTGTCACAAAGAGACATGCCCCCATGCGCTCCCTGTGATTTCCACGAGC	1140
QY	2408	TGCGCTCGGAGCGCTCCCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	2467
DB	1141	TGCGCTCGGAGCGCTCCCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	1200
QY	2468	AGGCTACAGCAGTACAGTACGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	2527
DB	1201	AGGCTACAGCAGTACAGTACGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	1260
QY	2528	TGGAGGACAGCGCGCCAGCTGAGGATATAGACCAATCAAGCAACATCTCAGCAGCAG	2587
DB	1261	TGGAGGACAGCGCGCCAGCTGAGGATATAGACCAATCAAGCAACATCTCAGCAGCAG	1320
QY	2588	GTCCCAACCATGGGCTGAACCTTGGAGAGACCTGGAGAGCTGCCCGCCCAAAAGTCCAC	2647
DB	1321	GTCCCAACCATGGGCTGAACCTTGGAGAGACCTGGAGAGCTGCCCGCCCAAAAGTCCAC	1380
QY	2648	AGCGGAGCGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2707
DB	1381	AGCGGAGCGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1440
QY	2708	GGCTGGAATGACAGCAGCTCCCTTCTACGCGGTGACATATAGAGAGAGCTTACCCACGA	2767
DB	1441	GGCTGGAATGACAGCAGCTCCCTTCTACGCGGTGACATATAGAGAGAGCTTACCCACGA	1500
QY	2768	ACTGCTCAGCAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	2827
DB	1501	ACTGCTCAGCAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1560
QY	2828	ATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2887
DB	1561	ATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1620
QY	2888	CCAGAGAGTGTGACTCCATCAGTGTGACAGCTCCAGAGAGAGAGAGAGAGAGAGAGAG	2947
DB	1621	CCAGAGAGTGTGACTCCATCAGTGTGACAGCTCCAGAGAGAGAGAGAGAGAGAGAGAG	1686
QY	2948	TCCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	3007
DB	1681	TCCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1740
QY	3008	CGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3067
DB	1741	CGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1800
QY	3068	TGAAGCCCAATGATGCTGTGATCAATAA	3093
DB	1801	TGAAGCCCAATGATGCTGTGATCAATAA	1825
REFERENCE	AK027471 Homo sapiens cDNA FLJ14565 fis, clone NT2RM4000233, highly similar to Mus musculus semaphorin 7A mRNA.		
AUTHORS	Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Wagatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,		

	Sugawara M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Wakamatsu, S., Kimura, K., Murakami, K., Ishii, S., Kawaji, T., Saito, K., Yamano, O., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuhara, Y., Nishimura, K. and Iwawata, T.
TITLE	NEDO human cDNA sequencing project.
REFERENCE	2 (bases 1 to 2293)
AUTHORS	Isogai, T. and Otsuki, T.
TITLE	Direct Submission
JOURNAL	Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1533-3 Yana, Kiseizaku, Chiba 252-0812, Japan (E-mail: genomics@helix.co.jp, Tel: 81-438-52-1975, Fax: 81-438-52-3886)
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan. cDNA full insert construction. Research Association for Biotechnology: cDNA library construction. 5' - 3' end one pass sequencing and clone selection. Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

[illegible]

```

Db      138: AGCGGGAGGCTCCCTGGGTCCCGGGAGGCCCTCCCTGCTCAGACCGCTCTAAGCAGGC 1440
QY      2708 GCGTGAATATCCACACCTCTCTTCCAGCGGTTTGACTATAAGAGGAGTACCCACGCA 2767
Db      1441 GCGTGAATATCCACACCTCTCTTCCAGCGGTTTGACTATAAGAGGAGTACCCACGCA 1500
QY      2768 ACTCGTCTACGAGAGCCACAGGCTACCACTCTCAAAAGAAACACACTAACTCTCTCA 2827
Db      1501 ACTCGTCTACGAGAGCCACAGGCTACCACTCTCAAAAGAAACACACTAACTCTCTCA 1560
QY      2828 ATTCTCTCTACCTCTCTCAGNAACAGAGCTTTGGCAGGGAGACACCGCCGCCGCC 2887
Db      1561 ATTCTCTCTACCTCTCTCAGNAACAGAGCTTTGGCAGGGAGACACCGCCGCCGCC 1620
QY      2888 CGCAGAGGGTGGAGCTCCATCCAGGTCCACAGCTCCCGACCATCTGGCCAGGGCGTCACTG 2947
Db      1621 CGCAGAGGGTGGAGCTCCATCCAGGTCCACAGCTCCCGACCATCTGGCCAGGGCGTCACTG 1680
QY      2948 TCTCGAGGAGCCGAGGCTCAAGCGCTACAACTCACTGACAAAGTGGGGGCTGAAGCGTA 3007
Db      1691 TCTCGAGGAGCCGAGGCTCAAGCGCTACAACTCACTGACAAAGTGGGGGCTGAAGCGTA 1740
QY      3008 CGCCCTCGCTTAAGCGGAGGTACCCGCCCAAGCACTCTTCTCGCCCTTTCAGATCCA 3067
Db      1741 CGCCCTCGCTTAAGCGGAGGTACCCGCCCAAGCACTCTTCTCGCCCTTTCAGATCCA 1800
QY      3068 TGAAGCCCAATGATGCTGTACATAA 3093
Db      1801 TGAAGCCCAATGATGCTGTACATAA 1826

RESULT 14
RD127394
LOCUS      2227 bp     DNA     linear     PAD 18-SEP-2002
DEFINITION      Primer for synthesizing full-length cDNA and use thereof.
ACCESSION      RD127394
VERSION      RD127394.1 GI:23222339
KEYWORDS      JP 2002017375-A/2825.
SOURCE      Homo sapiens (human);
ORGANISM      Homo sapiens
REFERENCE
AUTHORS      Bukaryola; Metazora; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases to 2227);
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakatsuki,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE      Primer for synthesizing full-length cDNA and use thereof
JOURNAL      Patent: JP 2002017375-A 2825 22-JAN-2002;
COMMENT
CS      Homo sapiens (human)
PN      JP 2002017375-A/2825
PD      22-JAN-2002
PF      07-JUL-2000 JP 2300253172
PI      TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI      ISHII,
PI      YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHIN-CHI KOJIMA,
PI      TETSUJI OTSUKI,HEISASHI KOGA
PC      C12N15/09,C07K14/47,C07K16/28,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
10,C12N1/02,C12N1/58//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
PC      Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT      CDS      Location/Qualifiers
1..2227
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT      567 a 700 c 560 g 400 t
ORIGIN

```

```

Query Match      56.3%; Score 1742.4; DB 6; Length 2227;
Best Local Similarity 99.9%; Pred No G;
Matches 1743; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1350 TTGSCCAGAAATAGGAATAFAUTGSGTTTCTTAATACAGAGCCCTTTCTCTGGAGGAGATGAG 1409
Db      1 TTGSCCAGAAATAGGAATAATGTTGTTTCTTAATGACAGCCCTTTCTCTGGAGGAGATGAG 60
QY      1410 TGTTTACACACTCTGAAAATAAGCAGCTATGATGGATCGAAGACAAAAGATCATGGGAT 1469
Db      61 TGTTTACACACTCTGAAAATAAGCAGCTATGATGGATCGAAGACAAAAGATCATGGGAT 120
QY      1470 CGAGCTGGACAGACGAAGCAGCTCTGTATGTTGGTCTCTACCTGTGTGATAAAGGT 1529
Db      121 CGAGCTGGACAGACGAAGCAGCTCTGTATGTTGGTCTCTACCTGTGTGATAAAGGT 180
QY      1530 TCCCTTGGCCGGTGTGAACGACATGGGAAGTGTAAAAAAGCTGTATTGGCTCCAGAGA 1589
Db      181 TCCCTTGGCCGGTGTGAACGACATGGGAAGTGTAAAAAAGCTGTATTGGCTCCAGAGA 240
QY      1590 CCATATGTTGATGATAAAGGAAGGTGGTGGCTGCAGCCATTTATACCCCAACAGCAG 1649
Db      241 CCATATGTTGATGATAAAGGAAGGTGGTGGCTGCAGCCATTTATACCCCAACAGCAG 300
QY      1650 ACTGACTTTTGAGCAGGACATAGAGCGTGGCAATACAGATGCTCTGGGGAGCTGTCAAA 1709
Db      301 ACTGACTTTTGAGCAGGACATAGAGCGTGGCAATACAGATGCTCTGGGGAGCTGTCAAA 360
QY      1710 TTTCTTTGGGACCTGCAATGGCATTCAGTTTCCTCTTTGCCAGCACAACACATCAGA 1769
Db      361 TTTCTTTGGGACCTGCAATGGCATTCAGTTTCCTCTTTGCCAGCACAACACATCAGA 420
QY      1770 TTTCCAGGCTCAAGAGGGATGAGTGTAGGAGGAGAAATGCTGGACTGGAAGCATCTGCT 1829
Db      421 TTTCCAGGCTCAAGAGGGATGAGTGTAGGAGGAGAAATGCTGGACTGGAAGCATCTGCT 480
QY      1830 TGACTCACTGACAGCAGACAGACCTTTGGGGCAGTGTCTTCCCATATATCAACAAGAAA 1889
Db      481 TGACTCACTGACAGCAGACAGACCTTTGGGGCAGTGTCTTCCCATATATCAACAAGAAA 540
QY      1890 GAAGGAGTGAATTCGGGAAGCTTACCTCAAGAGGACAGCAGCTGGTTCGGCTCAACCT 1949
Db      541 GAAGGAGTGAATTCGGGAAGCTTACCTCAAGAGGACAGCAGCTGGTTCGGCTCAACCT 600
QY      1950 CTTCGCCCTTCAGTCACTCTGGCTTCTCTATGAGGAGCGGCTCTCTCGAGCATCACCT 2009
Db      601 CTTCGCCCTTCAGTCACTCTGGCTTCTCTATGAGGAGCGGCTCTCTCGAGCATCACCT 660
QY      2010 CTACTGGCTCTGTGATCATCGCGCAAGACGCTGGCTGTGTGTCAGCGCAAGGAGAAGA 2069
Db      661 CTACTGGCTCTGTGATCATCGCGCAAGACGCTGGCTGTGTGTCAGCGCAAGGAGAAGA 720
QY      2070 GCTCAGCCACTCGCGCCGGGCTCCATGAGCAGCGTCCACAGCTCAGCGGCTCTTTGG 2129
Db      721 GCTCAGCCACTCGCGCCGGGCTCCATGAGCAGCGTCCACAGCTCAGCGGCTCTTTGG 780
QY      2130 GGACACTCAATCCAAAGACCCCAAGCGGAGCCATCTCTACGCCACTCATGCAACGCG 2189
Db      781 GGACACTCAATCCAAAGACCCCAAGCGGAGCCATCTCTACGCCACTCATGCAACGCG 840
QY      2190 CAAGTTCGCCACTTCGCCCAACACAGCGGCAAGATGCTCATTAAGCAGACACGACACCT 2249
Db      841 CAAGTTCGCCACTTCGCCCAACACAGCGGCAAGATGCTCATTAAGCAGACACGACACCT 900
QY      2250 GAGCTGAGCGGCTTCGCCCAACCCCAAGAGTCAACCCCAACCGCTGCGCAGAGAGCGGAAGCC 2309
Db      901 GAGCTGAGCGGCTTCGCCCAACCCCAAGAGTCAACCCCAACCGCTGCGCAGAGAGCGGAAGCC 960
QY      2310 CAGCGCGGCGAGCGCGGAGTGGGAGAGAACCAAGACCTCATCATGCTGCACAAAGGA 2369
Db      961 CAGCGCGGCGAGCGCGGAGTGGGAGAGAACCAAGACCTCATCATGCTGCACAAAGGA 1020
QY      2370 CATGCCGCCCATGGGCTCCCTGTGATTCCCAAGGACCTGCTCCCTGGGGGCTTCCCCGAG 2429

```

```

Db      1021  CATGGGCGCATGGGCTCCCTGTGATTCACAGACCTGGGCTGGGGCTCCGGAG 2082
QY      2430  CACATCCCGACGCGTGGTGGTCCCATCAGCAGCAGACCTACCAAGTACAGT 2439
Db      1081  CCATATCCCGACGCGTGGTGGTCCCATCAGCAGCAGACCTACCAAGTACAGT 1140
QY      2490  GCACGACGCCAANTGACGAGGTGGCCAGTGGCTGAGACACCAAGCGCCAGCT 2549
Db      1141  GGACGAGGCCAANTGACGAGGTGGCCAGTGGCTGAGACACCAAGCGCCAGCT 1200
QY      2550  GGATATAGACATCAAGACACATCTGACGAGGTGGCCAGTGGCTGAGACACCA 2609
Db      1201  GGATATAGACATCAAGACACATCTGACGAGGTGGCCAGTGGCTGAGACACCA 1260
QY      2610  TGTGAGACCTGACAGCGCTGGGCGGCTAAAGTTCACAGAGGAGAGCGCTCC 2669
Db      1261  TGTGAGACCTGACAGCGCTGGGCGGCTAAAGTTCACAGAGGAGAGCGCTCC 1320
QY      2670  CCGGAGGCGCTCCGCTGACAGCGCTGTAAGAGGAGGTGTAAGTCCACCACT 2729
Db      1321  CCGGAGGCGCTCCGCTGACAGCGCTGTAAGAGGAGGTGTAAGTCCACCACT 1380
QY      2730  TCCCTACGGGTGACTATAGAGAGGTACCGCAGGAGTGGCTGACAGAGAGCA 2789
Db      1381  TCCCTACGGGTGACTATAGAGAGGTACCGCAGGAGTGGCTGACAGAGAGCA 1440
QY      2790  GCGCAGGAGTCTGTAAGAGAGTACCTGCTGCAATCTGCTGCTGCTGCTGCA 2849
Db      1441  GCGCAGGAGTCTGTAAGAGAGTACCTGCTGCAATCTGCTGCTGCTGCTGCA 1500
QY      2850  CCAGAGCTTTGCGAGGAGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2909
Db      1501  CCAGAGCTTTGCGAGGAGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1560
QY      2910  GTCGACAGCTCCGAGCGCTGCGCGAGCGCTGCTGCGAGGAGCGCGCGCGCG 2969
Db      1561  GTCGACAGCTCCGAGCGCTGCGCGAGCGCTGCTGCGAGGAGCGCGCGCGCG 1620
QY      2970  CGCTACAGCTCAGTACAGAGCGTGGCGCTGAGCGCTGAGCGCTGAGCGCTG 3029
Db      1621  CGCTACAGCTCAGTACAGAGCGTGGCGCTGAGCGCTGAGCGCTGAGCGCTG 1680
QY      3030  ACCCGCCCAACCATCTTGTGCTGCTTCCACATCCATGAGGCGCAATGCTGT 3089
Db      1681  ACCCGCCCAACCATCTTGTGCTGCTTCCACATCCATGAGGCGCAATGCTGT 1740
QY      3090  ATAA 3093
Db      1741  ATAA 1744

RESULT 15
LOCUS   AK074975                2227 bp     mRNA     linear     PER 03-SEP-2002
DEFINITION Homo sapiens cDNA FLJ53494 fis, clone NT28P003614.
ACCESSION AK074975
VERSION   AK074975.1 GI:22760763
KEYWORDS  oligo capping; fis (full insert sequence).
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1
AUTHORS  Tsouai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
  Sugiyama,T., Suzuki,Y., Nagai,K., Suezo,S., Ishii,S.,
  Kawai-Hio,Y., Saito,K., Yamamoto,I., Wakamatsu,A., Nakamura,Y.,
  Kojima,S., Nagahara,K., Masuda,Y., Chono,K., Yoshikawa,Y.,
  Aotsuka,S., Sasaki,N., Hatlori,A., Okumura,K., Iwayanagi,I. and
  Nishimura,K.
  NEDO human cDNA sequencing project.
  Unpublished
  2 (bases 1 to 2227);

```

```

Tsouai,T. and Otsuki,T.
Direct Submission
Submitted (25-MAR-2002) Takao Isoqai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection; Helix Research Institute (supported
by Japan Key Technology Center etc.).
Location/Qualifiers
1..2227
/organism="Homo sapiens"
/seq_id="NT28P003614"
/clone="NT28P003614"
/cell_line="NT2"
/cell_type="retalorcarcinoma"
/clone_lib="NT28P3"
/notes="cloning vector: pXELSF3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction"
55..1744
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAC11326.1"
/db_xref="GI:22760763"
/translation="MSVNSEKSYGVEDKXIMQLDRSSSYVAFTCVTKVPL
GCEHGGKCTCIASRPYCGWIKEGGACSHLSPNSRLTTEQIERNTDGLGCHN
SEVATLHGSSLLPSTTSDTAQEGYSESGGNDLWKLSDPSTDLGAVSSHHQ
DKGVIRESYLGHPQLVLLATVILAFVNGAVFSGITVYVCHRRKDXAVVOR
KEKEITHSRSGMSVTVKLSGLFGDTQSKDPKFAILLPMHNGKLPFGNTAKMLIK
ADQHLIDTALPTPSTPTLOOKRPSRSEKERNONLINAC-KDMPGMSPLVPTD
LSPKSHVNLVWVLLPQVQREASUGPPGASLQGLSKEMHSSSTGVYKRR
SYNTSLTSKSHOATLKRNTSSNSHLRSQSGDSDPPAPQPVVDSTOVHSSQP
SOVATVSPRSPLNAYNSLRSGLKRTSLKPDVPPKPSFAPLSTSMKPNDACT"
567 a 700 c 560 g 400 t
BASIC COUNT
ORIGIN
Query Match 56.18; Score 1742.4; DB 9; Length 2227;
Best local Similarity 99.9%; Pred. No. 0;
Matches 1743; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1350 TTTCGGCAGAAATAGGAATAGTGGTTTCTAAATGACAGCCTTTCTCGGAGGAGATGAG 1409
DE 1 TTTCGGCAGAAATAGGAATAGTGGTTTCTAAATGACAGCCTTTCTCGGAGGAGATGAG 60
QY 1410 TTTTACAACTCTGAAAAATGCGAGCTATGATGAGTCCGAGGACAAAGGATCATGGCAT 1469
DE 61 TGTTCACAACTCTGAAAAATGCGAGCTATGATGAGTCCGAGGACAAAGGATCATGGCAT 120
QY 1470 GCAGCTGGACAGAGCAGGAGCTCTGTATATCTGCTCTACCTGCTGATGATGATGAT 1529
DE 121 GCAGCTGGACAGAGCAGGAGCTCTGTATATCTGCTCTACCTGCTGATGATGATGAT 180
QY 1530 TCCCTTGGCGCGTGTGCAACGACATGGGAAGTGTAAATAAAGCTGTATTTGCTCCAGAGA 1589
DE 181 TCCCTTGGCGCGTGTGCAACGACATGGGAAGTGTAAATAAAGCTGTATTTGCTCCAGAGA 240
QY 1590 CCATATTTGCAATGAGTAAAGAGCTGGTGGCTGCGAGCCATTTATCACCAACAGCAG 1649
DE 241 CCATATTTGCAATGAGTAAAGAGCTGGTGGCTGCGAGCCATTTATCACCAACAGCAG 300
QY 1650 ACTGACTTTGACAGCAGACATAGAGCTGGCATACAGATGGTCTGGGGAGCTGTCAACA 1709
DE 301 ACTGACTTTGACAGCAGCAGATAGAGCTGGCATACAGATGGTCTGGGGAGCTGTCAACA 360
QY 1710 TTCTTTGGCGACTGAATGGGCAITCCAGTTCCTCTTGGCCAGCACAACACATCAGA 1769
DE 361 TTCTTTGGCGACTGAATGGGCAITCCAGTTCCTCTTGGCCAGCACAACACATCAGA 420

```



QY	1770	TTCCAGCGGCTCAAGAGGGG:ATGAGTCTAGGGGAGGAATGCTGGACTGGAAGCACTGTGCT	1829
DB	421	TTCCAGCGGCTCAAGAGGGG:AACAGTCTAGGGGAGGAATGCTGGACTGGAAGCACTGTGCT	480
QY	1830	TGACTCAGCTGACAGACAGACAGCCCTTGGGGGCAATGTCTTCCCATTAATACACAAGACA	1889
DB	481	TGACTCAGCTGACAGACAGACAGCCCTTGGGGGCAATGTCTTCCCATTAATACACAAGACA	540
QY	1890	GAAGGAGTGATTGGGAAAGTTAGCTCAAGGCCACAGCCAGCTGGTTCGCTCAACCT	1949
DB	541	GAAGGAGTGATTGGGAAAGTTAGCTCAAGGCCACAGCCAGCTGGTTCGCTCAACCT	600
QY	1950	CTTGGGCATTGCAGTATCTTGCTTTCCTTATGAGGGGCGGCTTCTCGGGCATCAACCT	2009
DB	601	CTTGGGCATTGCAGTATCTTGCTTTCCTTATGAGGGGCGGCTTCTCGGGCATCAACCT	660
QY	2010	CTACTCGGCTGTGTATCATGCGCGGCAAGAGCTGAGCTGTGTCACAGCGCAAGAGAGGA	2069
DB	661	CTACTCGGCTGTGTATCATGCGCGGCAAGAGCTGAGCTGTGTCACAGCGCAAGAGAGGA	720
QY	2070	GCTCACCCACTCGCGGGCGGCGCTCCATGACAGCAGCTCAACAACTCAGCGGGCTCTTGG	2129
DB	721	GCTCACCCACTCGCGGGCGGCGCTCCATGACAGCAGCTCAACAACTCAGCGGGCTCTTGG	780
QY	2130	GGACACTCAATCAAAAGACCTAAAGGCGGAGGCCATGCTTATGCCACTCAATGACAAAGG	2189
DB	781	GGACACTCAATCAAAAGACCTAAAGGCGGAGGCCATGCTTATGCCACTCAATGACAAAGG	840
QY	2190	CMACTCGGCACCTCGCGGCAAGAGCGGCAAGCTGATTAATTAAGCAGCTCAGACCACT	2249
DB	841	CMACTCGGCACCTCGCGGCAAGAGCGGCAAGCTGATTAATTAAGCAGCTCAGACCACT	900
QY	2250	GGACTTGACGGGCGCTGTACACCCACAGAGTCAGAGTCAAGCTTCAC:ATAAGCGGAGGC	2309
DB	901	GGACTTGACGGGCGCTGTACACCCACAGAGTCAGAGTCAAGCTTCAC:ATAAGCGGAGGC	960
QY	2310	CAGCGCGGCTAGGCGGAGTGGAGAGTAAACACAGAGCTGATTAAGCTGACAAAGGA	2369
DB	961	CAGCGCGGCTAGGCGGAGTGGAGAGTAAACACAGAGCTGATTAAGCTGACAAAGGA	1029
QY	2370	CATGCCCGCATGGCTCCGCTGTGATTCACAGGAGTGCAGCTTGGGGCGCTCCCGGAG	2429
DB	1021	CATGCCCGCATGGCTCCGCTGTGATTCACAGGAGTGCAGCTTGGGGCGCTCCCGGAG	1080
QY	2430	CCACATCCGAGAGTGGTGATTCCTGGCCATCAGAGCAAGAGTATAGCATAGTACCT	2489
DB	1081	CCACATCCGAGAGTGGTGATTCCTGGCCATCAGAGCAAGAGTATAGCATAGTACCT	1140
QY	2490	GGACACCGCAAAATGAGCTAGTTCGCGCAAGATGGCTGTGGAGATAGGCTGGGACACT	2549
DB	1141	GGACACCGCAAAATGAGCTAGTTCGCGCAAGATGGCTGTGGAGATAGGCTGGGACACT	1200
QY	2550	GGAGTATAAGACCATCAAGCAAGCATTCAGCAGCAAGAGTTCGCAACATGGTGTGACCT	2609
DB	1201	GGAGTATAAGACCATCAAGCAAGCATTCAGCAGCAAGAGTTCGCAACATGGTGTGACCT	1260
QY	2610	TGTGGAGAACTGTGACAGCCTGGCCGCCAAGTTCCACAGCGGAGGCGCTCCCTGGGTC	2669
DB	1261	TGTGGAGAACTGTGACAGCCTGGCCGCCAAGTTCCACAGCGGAGGCGCTCCCTGGGTC	1320
QY	2670	CCGGGAGGCTCCCTGTCTCTCAGAGCGGTCTAAGCAGCGGCTGGAATGTACCACTCTCT	2729
DB	1321	CCGGGAGGCTCCCTGTCTCTCAGAGCGGTCTAAGCAGCGGCTGGAATGTACCACTCTCT	1380
QY	2730	TTCTACGGGCTTGACTATAGAGGAGCTACCCCAACAACTGCTCAGGAGAGGCCACCA	2789
DB	1381	TTCTACGGGCTTGACTATAGAGGAGCTACCCCAACAACTGCTCAGGAGAGGCCACCA	1440
QY	2790	GGCCACCACCTCTCAAAAGAAACACACTACTGTGTCAATTCCTTCTCACTCTCCACAAA	2849
DB	1441	GGCCACCACCTCTCTCAAAAGAAACACACTACTGTGTCAATTCCTTCTCACTCTCCACAAA	1500

[illegible]

Search completed: September 30, 2003, 20:16:09  
Job time : 11088.7 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.  
US-09-856-681-2  
October 23, 2003, 17:29:07 : Search time 26 Seconds  
(without alignments)  
3809.759 Million cell updates/sec  
US-09-856-681-2  
effect score: 5450  
sequence: 1 MRSEALLVFTLLHPAGAG.....PKKSEFAP-STSKXNDACT 1030  
coring table: BLOSUM62  
Gapop 10.0 , Gapex: 0.5  
searched: 283308 seqs, 96168682 residues  
total number of hits satisfying chosen parameters: 283308  
minimum DB seq length: 0  
maximum DB seq length: 200000000  
post-processing: Minimum Varch 0%  
Maximum Varch 100%  
Listing first 45 summaries  
Database : PIR 761\*  
1: PIR1\*  
2: PIR2\*  
3: PIR3\*  
4: PIR4\*  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

result	No.	Score	Query Match	Length	DB ID	Description
1	971.5	17.8	710	2	CH0798	fasciclin IV precursor
2	875	16.1	656	2	B49423	semaphorin III - fru
3	852.5	15.6	711	2	D49423	semaphorin III - fru
4	845.5	15.5	711	2	A49423	semaphorin III - fru
5	842.5	15.5	712	2	A49423	semaphorin III - fru
6	839.5	15.4	712	2	A49423	semaphorin III - fru
7	830	15.2	704	2	C03228	collapsin - chick
8	826	15.2	749	2	G01856	semaphorin F - hum
9	803	14.7	748	2	I48744	semaphorin A - mol
10	793.5	14.6	666	2	I58169	semaphorin III - m
11	789	14.5	712	2	I27165	hypothetical prote
12	779.5	14.3	724	2	C49423	semaphorin II - fru
13	768	14.1	753	2	G02273	semaphorin III - fru
14	737	13.5	751	2	I48748	semaphorin E - mol
15	692	12.7	814	2	S66498	X-sema F protein p
16	656	12.0	782	2	I48746	semaphorin C - mol
17	641.5	11.8	760	2	I48745	semaphorin B - mol
18	356.5	6.5	653	2	T03132	semaphorin homolog
19	322.5	5.9	676	2	T33833	hypothetical prote
20	238	4.4	1945	2	T13337	Plexin A - fruit f
21	217	4.0	403	2	A42521	A398 protein - vac
22	205	3.8	441	2	S29921	hypothetical prote
23	199.5	3.7	1884	2	C49475	Plexin 2 precursor
24	182	3.3	1905	2	T51552	Plexin - African c
25	163.5	3.0	3968	2	A44265	trifolax homolog
26	159	2.9	2051	2	T13164	Plexin B - fruit f
27	155.5	2.9	1894	2	C49480	Plexin 1 precursor
28	155	2.8	1375	2	C05146	hepatocyte growth
29	155	2.8	2352	2	T30201	Notch homolog Proc

30 153 2.9 625 2 S48941 regulatory protein  
31 147.5 2.7 3507 2 T34513 hypothetical prote  
32 147 2.7 1425 2 T30811 hepatocyte growth  
33 146.5 2.7 295 2 Q01775 Sal19R protein - v  
34 140 2.6 1390 1 TVHUME hepatocyte growth  
35 139.5 2.6 3869 2 A48205 All-1 protein - GTE  
36 138.5 2.5 1065 2 S19482 hypothetical prote  
37 138 2.5 728 2 S48569 hypothetical prote  
38 137 2.5 867 2 T41308 hypothetical zinc-  
39 137 2.5 2492 1 C44213 nonstructural poly  
40 136.5 2.5 1347 2 T02214 ubiquitous TPR mot  
41 136 2.5 1310 2 T43135 exysterc.-binding  
42 136 2.5 2531 2 S18188 notch protein: homo  
43 133.5 2.4 1829 2 T14280 RMI protein - mod  
44 133.5 2.4 2187 2 T30826 nascent polypeptid  
45 133.5 2.4 2531 2 A46019 notch-1 protein -

ALIGNMENTS

RESULT :  
CH0798  
fasciclin IV precursor - American bird grasshopper  
C:Species: Schistocerca americana (American bird grasshopper)  
C:Date: 11-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 27-May-1999  
C:Accession: JH0798  
R:Kolodkin, A.; Matthes, D.C.; O'Connor, T.P.; Pate, N.H.; Admon, A.; Bentley, D.; See  
Neuron 9, 831-845, 1992  
A>Title: Fasciclin IV: Sequence, expression, and function during growth cone guidance  
A:Reference number: JH0798; MUID:93040225; PMID:1418998  
A:Accession: JH0798  
A:Molecule type: mRNA  
A:Residues: 1-710 <KOL>  
A:Cross-references: GB:J00703; NID:G160944; PID:G160945  
A:Experimental source: embryo  
A:Comment: This protein plays a role in growth cone guidance in the developing central  
C:Keywords: glycoprotein; trans-membrane protein  
F123-710/Domain: signal sequence #status predicted <SIG>  
F123-710/Product: fasciclin IV #status predicted <KAT>  
F123-627/Domain: extracellular #status predicted <EXT>  
F123-652/Domain: transmembrane #status predicted <TM>  
F1653-710/Domain: intracellular #status predicted <INT>  
F144-711/63-267,360,539/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	17.8%	Score	971.5	DP	2	Length	730
Best Local Similarity	34.4%	Prod. NO.	5.4e-57				
Matches	244	Conservative	124	Mismatches	224	Indels	117
Gaps	29						
QY	5	LYFTLLHPAGGFFDELSEPISSHGNTYKQYVPVGHKPSRNTTCR					
DB	11	LIMVALCH--AAAWNVKVP-----KMYVQ-----SEYVQRFUGNESHDKHFKLJ					
QY	62	NIMMGSLVIAARDHYTVQID--TSHTEEIYCSKKLTWKSQCADVTCRYKGRKDECHN					
DB	55	EKHNSLLVGARNVYNISJRELTEPTEQ-----RIEWHSSGAHRELCYLGKSEDDCCN					
QY	120	FIKVLKKNDALFVCGTNAPNSCRNYKMDTLEFFGD-----EFSGMARCPYDAKHANV					
DB	110	VIRVLAKIDDRVLCGGINAYKPLCRHYALKD---GDYVVEKEVEGRGLCPDPDHNST					
QY	175	ALPAGKLYSATVTEFLAIDAVIYRISGESPTLTIVHDSKNWKEPFEQVADVGDVIYF					
DB	166	AIYSEGQLYSATVADFSGTDFLIYRG-----PLRTSRDLKQLNAPNFVNTMYNDPIFF					
QY	235	FFREIAYEYNTMGVVFPRVACVCKNDMGSSORVLEKWTSEFLKRLNCSVPGDSHYFEN					
DB	221	FFREIAYEYNTMGVVFPRVACVCKNDMGSSORVLEKWTSEFLKRLNCSVPGDSHYFEN					
QY	235	ILQAVTVIRIN-GRDV---VLATFTSPYNSIPGSAVCAYDMLDJIASVFTGRPEKQSPD					
DB	230	BIQSTSDIIEGNYGGQVEKLIYGVFTFTVNSIGSAVCAFSMKSLLESFGGPFPEQETMN					





b 485 --BPTTISAVELSTKQQLVIGSTAGVAGQPLAKKCDIYKACACCECIAREPVCAM--DGS 540  
y 541 ACSHLSPKS-PLTFEQLTERGNTGGLDCHNSFVALNGHSSLL 542  
b 543 SCSSPYPTAKSARTRQPIKNG--DPLTHCSLDELHCHNHGKPSL 541  
RESULT 6  
49369  
collapsin - chicken  
;Species: Gallus gallus (chicken)  
;Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 24-Sep 1994  
;Accession: A49069  
;Uco, F.; Raible, D.; Raper, C.A.  
ell 75, 217-227, 1993  
;Title: Collapsin: a protein in brain that induces the collapse and paralysis of neurons  
;Reference number: A49069; PMID:8402654; PMID:8402654  
;Accession: A49069  
;Status: preliminary; not compared with conceptual translation  
;Molecule type: mRNA  
;Residues: 1-772 <Uco>  
;Cross-references: GI:506228; NID:34-0078; PDB:1AC0963B.1; PDB:1AC0963B  
;Superfamily: semaphorin  
Query Match 15.4%; Score 839.5; DB 2; Length 772;  
Best local similarity 35.2%; Pred. No. 41e-48;  
Matches 198; Conservative 87; Mismatches 220; Indels 57; Gaps 17;  
/ 68 LYAARDHYTVVDIDTSHTREIYCSKALTKSKQADVDCTRMKGR-KDECHNFIXVLK 126  
o 69 LYVAKDHIFSNL-----VNIKEYQKIWVSVSHSRDECKWAGKILRECANFIKVLK 123  
y 127 KNDALFVCGTNAPNSCR-----NYKMTLEPPGDEF--SCVACGYAKANVALPAD 179  
b 124 YNQHETACGTGAPHPMCYIEVSHPEQWIFRMEDSHFENGKGRSPYKLLTAGLLVD 183  
y 180 GKLYSAVTOTPLADAVIYRSLGSEPTLRIVKDSKWKLEPFYVQA-----VDYGVY 232  
b 184 GELYSGLAADFVGRDPAFRTLGHHPHPIRTECHDSRWLNDPRISALIPESDNFEDDK 243  
y 233 YFPRETAVENTYKGVVFRVAOVKQNGGSGRVLKQWTSFLKARLNCSPG----D 288  
b 244 YFFRENAIDGHTGKATHARIGQCKNDFGG-HRSVYKWTTLKARLTCSEVGPWG-D 302  
y 289 SHYENILQAVTVIRNGSD---VVLATFSTPYNISPGSAVCAVDMLDIASVFTGRK 344  
b 303 THF-----DELGVFVNSKQPKPIVGYFTSSNIFKGSVCAVYKSTVRRVPLGPA 357  
y 345 EQKSPSTWTFVDERVYKPRPGCCAGSSSLERYATSNFPPDTLNFIFKTHP-LYDRAVPS 404  
b 358 HRDGPYQWVPY-QGRVPYPRPGTGP-SKTFGGFDSTKCLPDEVITPAKSRPMYKVEFP 415  
y 405 FNEPWLRTNWRVRLTKAVDTAAGPYQCHTVYVFGSEKGIILKPLARIGNSGFLNDSJ 464  
b 416 INSRFVTKIDVYCFCTQVVDVRVADGGQIDWPFQTDIGTVLKVYSVTKETWHIEEV 475  
y 465 FLEMSVYNSEKCSYGVDEKRIKNGMLDPAASSLVVAFSTQVIVKVLQRCEREGKCKYT 524  
b 476 LLEBYTFER-----EPTVLSAKYISTKQQLVIGSATGVSQFLKCDIYKACAE 526  
y 525 CIARDPYCWMKEGGACSHLSPNS-RLTTEQIERGNTDGLDCHNSFVALNGHSSLL 583  
b 527 CCLARDPYCAM--DGSSCSRYFPTAKRTRRQPIKNG--DPLTHCSLDELHCHNHG 575  
y 584 PSTTSDSTAQGEYVESGGMLD 605  
b 576 PSQVLEBKIIYGVKSSSTFLS 597

RESULT 7  
C5928  
semaphorin F precursor - human  
;Species: Homo sapiens (man)

CDate: 10-Apr-1998 #sequence\_revision 08-May-1998 #text\_change 17-Nov-2000  
CAccession: J05928  
R.Stammers, A.D.; Pleschel, A.W.; McPherson, C.D.; Overhauser, J.; Lovett, M.  
Biochem. Biophys. Res. Commun. 242, 665-691, 1998  
A>Title: Molecular cloning and mapping of human semaphorin F from the Cri-du-chat candi-  
AReference number: C5928; NID:38-2554; PMID:9464278  
AAccession: J05928  
A>Status: nucleic acid sequence not shown  
A.Molecule type: mRNA  
A.Residues: 1-1074 <SIM>  
A.Cross-references: GI:52840; NID:3272583; PDB:1AC09473.1; PDB:1G272584  
A.Experimental source: brain  
A.Comment: This protein disrupts normal brain development and leads to some of the feat-  
CGenetics:  
A.Gene: sema4  
CSuperfamily: human semaphorin F; thrombospondin type 1 repeat homology  
F1-10/Domain: signal sequence #status predicted <SIG>  
F10-53/Domain: semaphorin #status predicted <SE>  
F140-496/Domain: thrombospondin type 1 repeat homology <THR3>  
F171-993/Domain: transmembrane #status predicted <TM>  
Query Match 15.2%; Score 930; DB 2; Length 1074;  
Best local similarity 36.2%; Pred. No. 3e-47;  
Matches 199; Conservative 90; Mismatches 211; Indels 50; Gaps 17;  
QY 68 LYAARDHYTVVDIDTSHTREIYCSKALTKSKQADVDCTRMKGR-KDECHNFIXVLK 127  
DB 70 LYVAKDYIFRLQJ-----ECLSLQAVNECCBATKACYSKSKSECCQYIRVLV-V 123  
QY 128 NDDALFVCGTNAPNSCRNYKMTLEPPGDEFGMARCPYDAKFAVALF-ADGKIYSAT 186  
DB 124 GGDLFTCGTNAFTVCTNRSLSNLAIEHQISGMARCPYSPHNSFTALLTAGELYAAT 193  
QY 187 VTDLAIDAVIYRSLGSEPTLRIVKDSKWKLEPFYVQAVDYQDYIFFPRE-AVENTW 246  
DB 184 AMDFGDEPAIYRSLGILPLRTAQNSKWLNEPNFVSSYDIGNFYFFPRENAVEED-C 242  
QY 247 GKVVFPRVAOVKQNGGSGRVLKQWTSFLKARLNCSPVGDSHFVFNILQAVTVDIRIN 306  
DB 243 GKTVPRAARVCKNDTGG-RFLLEDTWTFMKRLNCSRGGEVFPFYINBLQSTFFLPEL- 300  
QY 307 GRVYLATFSTPYNISPGSAVCAVDMLDIASVFTGRFKQKSPDSTWTFVDERVPKPRP 366  
DB 301 -DLIYI-GFTVNSIIAASAVCVFNLSA-ACATSGFPKYQENSRSALPYEN--PNPHF 355  
QY 367 CCCAGSSSLERYATSNFPPDTLNFIFKTHPLMDEAVPSIPNRFPLATWVRYELTIAVD 426  
DB 356 QCGTVDDGLYVNLTERNL-QDAOKFLVH----EVVQPVTVVPSFMEDNSRP--SHVAVD 438  
QY 427 TAAGPYQWHTVYVFGSEKGIILKPLARIGNSGFLNDSLF-EEMSVMYNSKCSYDGVDECKR 486  
DB 409 VVQGPENLAVHIIYATDYGTKKVRVPLNQT---SSCCLEELIELEFPERR-----REP 458  
QY 487 IMGVLQDRASSLVVAFSTQVIVKVLQRCERHCKKKTCLIASRDPYCGWIKEGGASHLS 546  
DB 489 IRSQILHSCSVLFVGLPHEHWKIFLKRCQFY-RTSTCIGAQDPYCGMDVWKKCTSLB 517  
QY 547 PMSRLT-PEGDIERGNTDGLDCHNSFVALNGHSSLLPST--TSDSTAQGEYVESGGGM 603  
DB 518 ESLSMTQWESISA-----CPTNLTVDGFGVMSFWPCTHTDGSVAV-----GSC 563  
QY 604 LKMKHLLDSP 613  
DB 564 LCRTESCDSP 573

RESULT 8  
G01856  
semaphorin V - human  
C.Species: Homo sapiens (man)  
CDate: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 24-Sep-1999  
CAccession: G01856  
R.Sekido, Y.

Unsubmitted to the EMBL Data Library, June 1995  
;Reference number: G05634  
;Accession: G01856  
;Status: preliminary; translated from GB/EMBL/DBJ  
;Molecule type: mRNA  
;Residues: 1-749 <RES>  
;Cross-references: EMBL:U28369; NID:5974293; PID:AA09138.1; PID:G974254  
;Superfamily: semaphorin

Query Match 15.2%; Score 825; DB 2; Length 749;  
Best Local Similarity 34.7%; Pred. No. 3.2e-47;  
Matches 198; Conservative 89; Mismatches 226; Indels 56; Gaps 16;

```
Y 66 GTLYIAARDHIYTVCTDTSSTEEYCCKKLTNKSQADVDTCRYKQK-KDECHNFPKVL 124
D 65 GRLVGAENHVASLSDNISKR-----AKLWAPVVEWREECNAGKDICTEONKPFVRL 120
Y 125 LKNDALFVCGTNAFNSCRNRYKDE-----TLEPCEBSSGVARCPYDAKHANVA 176
D 121 HAYNRTHLACGTCAEPTCALRWATAGGTHASTGPEKLED--GKGTFFYPRHRPPS 177
Y 177 PADGKLYSATVTPFLAIDAVYRSJGESPTLRIVKHSKWLKRPYFQAV-----DY 229
D 180 LVGEELYSVAADLNGRDTIPRSLGQNPRLRPHDSRLNEPKFVKVFWIPSENP 239
Y 230 DYIYFFREFAVE-YNTMGKVPFVAQVCKNDGGSQVLEKQWTSFLKARLNCSPV- 286
D 240 DKYIYFFREFAVEAALGRLSVSRVQVCRNDVGG-QKSLVKNKWTTFUKARLVCSVP 294
Y 287 -GDSHYFENILQAVTVIRNGRD---VYLAFSTPYNSIPQSAYCAYDMLDIASVFT 341
D 299 EGTHF-----DGLQVFLSSRDHRTPLLYAVFSTSSIFQSSAVCVYSMNDVRAFL 353
Y 342 RFKEQKSPDSTWTFVDERVPRKPPCCAGSSSLRYATSNFPDPTLNFIKTHPLMDE 401
D 354 PFAKKEGEKQHWYSY-QGRVYVPRPGVCP-SKFTGTFSSSTKDFPDVIGRHHPLYN 411
Y 402 VPILPRPWFRLTMVRYLTKAVTAAGPYQHTVVLGSEKGIILKFLA---RIGN 456
D 412 VLPFGRLFLQVAGYTFPTQADRVAAADGHDYDLFIDTGVTVLKVSVPKGRPN 469
Y 462 DSLEFLEMSVYNSEKSYGVDEKRMGMQDIPAGSSLYVAESTCVIKVPLGRCERHGC 521
D 472 EGIILLDELHYFE-----DGAATSNQISKRHLVYASSASVAQAIALHRCAGHRY 522
Y 522 KKTCLASRDYPYCKWKEGGACSHLSPNSRLTF-EQDIERKNTDGGDCHNSFVALNGH- 578
D 523 CTECCUARPYPYCAW-EGVACTRFQSAZRRPRQVGRGDRSTLCSGDSRPALEHAY 580
Y 579 -----SSSLPSTITSDTAQGEYSEGGY 603
D 591 FGVEGSAFLECEPPSLQARVETFORAGV 610
```

## RESULT 9

48744  
semaphorin A - mouse  
;Species: Mus musculus (house mouse)  
;Date: 02-Jul-1996 #sequence\_revision: 02-Jul-1996 #text\_change 24-Sep-1999  
;Accession: I48744  
;Puschel, A.W.; Adams, R.H.; Betz, H.  
euron 14, 941-948, 1995  
;Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates  
;Reference number: I48744; NID:95267431; PMID:7748561  
;Accession: I48744  
;Status: preliminary; translated from GB/EMBL/DBJ  
;Molecule type: mRNA  
;Residues: 1-749 <RES>  
;Cross-references: EMBL:X85992; NID:5854323; PID:CAA599892.1; PID:G854324  
;Gene: sema  
;Superfamily: semaphorin

Query Match 14.7%; Score 803; DB 2; Length 748;  
Best Local Similarity 35.5%; Pred. No. 1.1e-45;  
Matches 191; Conservative 80; Mismatches 205; Indels 62; Gaps 18;

```
QY 66 GTLYIAARDHIYTVCTDTSSTEEYCCKKLTNKSQADVDTCRYKQK-KDECHNFPKVL 124
DB 65 GRLVGAENHVASLSDNISKR-----AKLWAPVVEWREECNAGKDICTEONKPFVRL 120
QY 125 LKNDALFVCGTNAFNSCRNRYKMDL-----EPFDEBSSGVARCPYDAKHANVA 175
DB 121 HAYNRTHLACGTCAEPTCALRWATAGGTHASTGPEKLED--GKGTFFYPRHRPPS 177
QY 176 PADGKLYSATVTPFLAIDAVYRSJGESPTLRIVKHSKWLKRPYFQAV-----DY 228
DB 178 VLVEELYSVGTADLNGRDTIPRSLGQNPRLRPHDSRLNEPKFVKVFWIPSENP 237
QY 229 DYIYFFREFAVE-YNTMGKVPFVAQVCKNDGGSQVLEKQWTSFLKARLNCSPV- 286
DB 238 DDKYIYFFREFAVEAALGRLSVSRVQVCRNDLGG-QKSLVKNKWTTFUKARLVCSVP 296
QY 287 -GDSHYFENILQAVTVIRNGRD---VYLAFSTPYNSIPQSAYCAYDMLDIASVFT 340
DB 297 VESDTHF-----DGLQVFLSSRDHRTPLLYAVFSTSSIFQSSAVCVYSMNDVRAFL 351
QY 341 RFKEQKSPDSTWTFVDERVPRKPPCCAGSSSLRYATSNFPDPTLNFIKTHPLMDE 400
DB 352 GFLPKKEGTHQWYSY-QGRVYVPRPGMCP-SKFTGTFSSSTKDFPDVIGRHHPLYN 409
QY 401 AVPSIERNRPFRLTMVRYLTKAVTAAGPYQHTVVLGSEKGIILKFLA---RIGN 456
DB 410 PVLPMGGRPLFLQVAGYTFPTQADRVAAADGHDYDLFIDTGVTVLKVSVPKGRPN 469
QY 457 SGFLNLSLFEEMSVYNSEKSYGVDEKRMGMQDIPAGSSLYVAESTCVIKVPLGRC 516
DB 470 S-----EGILLDELQVPE-----DGAATSNQISKRHLVYASSASVAQAIALHRC 516
QY 517 RHGCKTKTCASRDYPYCKWKEGGACSHLSPNSRLTF-EQDIERKNTDGL--GECHNS 571
DB 517 ALGRACACCLARPYPYCAW-DGSACTRFQTAZRRPRQVGRGDRSTLCSGDSHSHS 572

RESULT 10  
488169  
semaphorin II - mouse (fragment)  
;Species: Mus musculus (house mouse)  
;Date: 26-Jul-1996 #sequence_revision: 26-Jul-1996 #text_change 24-Sep-1999  
;Accession: I58169  
;R.Mosser-Smith, E.K.; Leonard, R.D.; Shatz, C.J.; Tessier-Lavigne, M.; Goodman, C.S.; K.  
Neuron 14, 949-959, 1995  
;Title: Semaphorin II: can function as a selective chemorepellent to pattern sensory p  
;Reference number: I58169; NID:95267432; PMID:7748562  
;Accession: I58169  
;A:Status: preliminary; translated from GB/EMBL/DBJ  
;Molecule type: mRNA  
;Residues: 1-666 <RES>  
;Cross-references: GB:I40484; NID:9703189; PID:AAA71934.1; PID:G703130  
;Gene: Sema11  
;Superfamily: semaphorin


```

Query Match 14.6%; Score 793.5; DB 2; Length 666;  
Best Local Similarity 34.1%; Pred. No. 4e-45;  
Matches 188; Conservative 89; Mismatches 226; Indels 49; Gaps 15;

```
QY 116 ECHNFIKVLKKNIDALFVCGTNAFNSCR-----NYKMDTLRPFQDEF--SGVARCPYD 168
DB 7 ECANFIKVLKYNQTHLYACGTGAPHPICITYIEVGHHPEDNIFKLQDSHFENRGKSPYD 66
QY 169 AKHANVALFADGKLYSATVTFDLAIDAVIYRSJGESPTLRIVKHSKWLKRPYFQAV--- 225
DB 67 PKLJATASLIDGELYSGTAANFMGRDFAIFRTLGHRRHPIRTEQHCERWLANDPRFISAE 126
QY 226 ----VDGVDIYFFREINVEYNMGKVPFVAQVCKNDGGSQVLEKQWTSFLKARL 281
```



```
150 TWANRP--KCVYIVANLTHLPSEVIGV---LGIACPYDPLDNLNIIYVENONPGGL 204
151
152 KLYSATVDFLAIDAVYRS-----LGSPTRTVKXDSKWLKTPYVQAVYDGY 231
153
154 POLYSGTNAEFTKATVAFRTDYNTSAKRLKYLKYLKYLKYLKYLKYLKYLKYLKYL 264
155
156 IYFFRETAIVEVNTGKVVPEPVAQVCKXVGGSGORVLEKQWTSFKARLNCSPGDSHP 291
157
158 VYFFRETAIVEVINGKAVYSIAIRVCKXVGG--KMLAHKWATYKARLNCSSISGEPP 323
159
160 YFNILCAVTDVIRINGROVVLATFTSPYNSIRGSVAVCYDMLCIAVSTIGRFKECKSPDS 351
161
162 YFNEOSVYQLPSDKSR--FFATFTTSTNGLGSAVCSFHINEICAAFGKRFKECKSSNS 381
163
164 TWTPVDERVPRGRCGCCGSSSLERYATSNFFPDITLNFIKTHLPXDEAVPSIFNRF-W 410
165
166 AWLPVJNSVPEPREGTCVNDTS-----NLPDVLNFIKTHLPXDKAVNHEHNPVY 433
167
168 PLRTVVRVRLT--KIAVDTAAGPYQKHTVFLGSEKGIILKFLASIGSGFNDLSFLER 469
169
170 YKRDVFTKLVDKLRIDL---NSYIVYVGTNGLRIYKLVQYRNGESLSKLLDIFE 490
171
172 MSVYSEKCSYGVDEKRIKNGVQLDPRASSLVVAFSTCVIKVFLGRC-ERRSKXKTCIA 527
173
174 VA-----PNEALQVMEISQTRKSLVGTGHRIKQDLANCKRRVYDNCPR-CV- 536
175
176 SDPYCGMIKCGAGSH:SPNSRLTPEQDIERGNTD 563
177
178 RCPYCGMDKEANTCRPY-----ELDLCQVANEATSD 567
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
```

```
277 CLVKK-WSTFKARLYCSVPGDS:ETHFDELQGVVQCTQDV---NPVIYVFTSSGS 332
321 SPGSAYCAYDMLCIAVFTGRFKECKSPDSTWTPVDERVPRGRCGCCGSSSLERYAT 380
333 VPRGSAYCVSVADISGVFNPGPTAHKEGPNYQWMPF--SGKMPYPRDTCGGTFTPMKS 391
381 SNFFPDITLNFIKTHLPXDEAVPSIFNRFKTHLPXDEAVPSIFNRFKTHLPXDEAV 437
392 TKDYPDGVIVNFMKSHPLXACAVYFLORRPLVAVTGAPYRLTTIAVQVCSADGRYE--V 448
438 VFLGSEKGIILKFLASIGSGFNDLSFLGSEKGIILKFLASIGSGFNDLSFLER 497
449 LFLGTGRGTQKIVV-LPKDOEVEELMLEVEVFK-----DPAPVKTWISSKQ 498
496 SJVAFSTCVIKVFLGRCERRGCKKTCIASRDPYCGWIKESGACSHLSPNS-RLTPEQD 536
499 QLYVASAVGVTHLSHRCCAYGAACADCC-LARDPYCAW---DGAQCSRYTASSKRRSRQD 556
557 IERN 561
557 VRHGN 561
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
```



```

D 368 GRTPYPAFGTCPPGGAFTPNVHTTQDFDQVVTFRNHPLMYNSISPIHRRPILVIGTGY 427
Y 419 RLTKIAVCTAAGPYQNHVTVFLGSEKGIILKFLARLNSQF.NKDSPLFEMSGVNSKES 478
D 428 KYKIADVDRUNADGRVHVFLGTDRCTQKVVVLTNS-ASGEILLELEVPKXHI--- 483
Y 479 YDGVEDKRLKMGMDLRASSSLYVAFSTCVIKVPLGRCERHGKAKTCTIASRDPYCGNIKE 538
D 484 -----VFITINEISSKKQGLYVSNFSGVSLHRCCHIYGTACACCCCLARDPYCAM--D 535
Y 539 GGACSHLSP.NSLTFPQDIERGHTDGLGDCNHSFVALAGHSSCLPSTTTSDSACE 595
D 536 GHSCSRYPYTKGARSARQVHGNA. LTPCRGPNLKATENALAEIVQYGVNNS.PLE 591

ESULT 15
66498
-sma F protein precursor - mouse
/Species: Mus musculus (house mouse)
/Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #taxi_change 24-Nov-1999
/Accession: S66498
/Inagaki, S.; Furuyama, T.; Iwahashi, Y.
EBS Lett. 370, 269-272, 1995
/Title: Identification of a member of mouse semaphorin family.
/Reference number: S66498; MUID:95384603; PMID:7456991
/Accession: S66498
/Status: preliminary
/Molecule type: mRNA
/Cross-references: EMBL:S70463; NID:G1110639; PDB:1AB154.1; PDB:G1110699
/Suprafamily: semaphorin
/1-21/Domain: signal sequence #status predicted <SIG>
/22-834/Product: N-terminus F protein #status predicted <NAT>

Query Match 12.78; Score 692; DB 2; Length 834;
Best Local Similarity 26.0%; Pred. No. 3.5e-38;
Matches 232; Conservative 131; Mismatches 126; Indels 204; Gaps 34;

/ 65 NGILYIARDHIVTDITDSTHETCYCKLTKWKSQAQVDTCKMKGKHKDCHNFIKY 123
D 62 SGLLYGAREALFAFSV-----EALQSAISWEAPAEKKIECTCKGNQTECFNFIFF 116
Y 124 LLKKNDALFVCGTNAFNPSCRNVMDTLEPPGDEF-SGVARCEYDARHANVALPADGKL 182
D 117 LOPNSSHLVCCCTYACPKCTYINMLPTLDRABEDGCKGKCPYPAKHTGLLVGEL 176
Y 183 YSATVTFDAIDAVIYRSLGESFTLRVTKSDKMLKEPYFVQA-----VOYGDVI 232
D 177 YSATLNNFLGTPEVILRYMGTHASIKT-EYAFNLKRPHPFVGSAPFESYGSFTGDDK1 235
Y 233 YEFREI-AVENTYGVYPRVAVQCKNDMGSCRYLEKOKTSFLAPRLNCSYFGDSHEY 232
D 236 YFFSERAVEYDCYSEGVAVARVAVCKGMGGA-RLQKKAITFLKARLVCSAP-ENKYY 293
Y 293 FNILQAVTDVIRINGROW-LATFTPYNSIPGAVCAVDMLDIASVFTSPFKEOKSPS 351
D 294 FNQKAVHLRAGSWHTTFGVFQARWGDYLSAVCEYQLECIQOVFSGPYKEYSEQAQ 353
Y 352 TWTVFPERVPERKRPCCGACSSSER-YATSNRPFDITLNFKTNFLNDEAVPSIFRPA 410
D 354 KWARYTDP-VESFAPGSCINNHFRDNGYTSLSLEPNTLAFIXHEFUMEDQVAF-GRFL 412
Y 411 PLRTWAVRLTKIAVDTAAG-PVGNHTVWFLGAEKGIILKFLARLNSQF.LNDSF--- 464
D 413 LVKXNTNF--THVTADRVPLDGLGTYVLTGTGCGNLKAV-----SLGPNWH 459
Y 465 PLEBNSVYNSKESYGVDEKXKINGMO.DRASSSLYVAFSTCVIKVPLGRCERHGKAKT 524
D 460 KWEELOWEDQEP-----VFSLVLSQSKXVIFAGSRSLQVLSLADCTWYRQC-VQ 564
Y 525 CIASRDPYCGNIKEGGAC-----SHLS--FNSRLTFEGDIER----- 569
D 509 CVLARDPYCAMNVTNTRCATTSRSGSFVGVHAKLDTSKXCNQYGIKVRSPIPKNTY 568

```

Search completed: October 23, 2003, 17:11:31  
 Job time : 30 secs

```

QY 560 -GNTGLGDCCHNSFVALNGH---SSLLRSTTTSDTAQEGYESRQGMLOWKHLLDSPS 615
D 569 VSGTDLVLPCHLSSNLAFAHWTGSCDLPAEQPGSF.YDTGLQALVW----- 616
Y 616 TDP-GAVSSN-----HQDKKG--VIRSEYKLG-----HCQLVPVTLIAIVILAFV 660
D 617 -----AASRHSRGPYCYSEBQGTRLAASLYLVAVAGSSVTLARAPLENLGLWKLAVYA 672
Y 661 MGAVFSSGTYVYCVCDHRRKDVAVVQKEXELTHSRRGSMSSVTKLSGLFGDTQSKDKPKE 720
D 673 LGAVCL-VILLVLVLSLRR-----LRSELE-----KGAKAS-----BR 704
Y 721 ALITPLVINGKLATPCTAKMKIQAQOHHLULTAFTPSTPTLQOKRKPSPKSGREN--- 737
D 705 TLVVPD-----ELPKEPASHPF--RPGPETCKLMDPV 735
Y 736 -----ERNQNLINACTKMDPPVGSPIPTDLPLRASPSHIPSWWV.PITCOGYCHEYVDQ 832
D 736 GYYYSDSGLKIVPGHARCCDGGGPPSPPPGIPGCPUPS--PTRLHLLGGSNNSNANGYVR- 792
Y 833 FKMSEVAQMALEDQAATLEYKTIKEHLSSKSPNHGYNLVENLDSLEPKYPORE 585
D 793 -----LQJGSEDRGGS-----GRPLFELATLRLKLOQRQ 822

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.  
M protein - protein search, using sw model.  
Run on: October 23, 2003, 17:09:07 : Search time 17 seconds  
(without alignments)  
2349.345 Million cell updates/sec  
Title: US-09-856-681-2  
Effect score: 5450  
Sequence: 1 MSEAALLVFTLLHPAGAGF.....PRKPSFAPLSMKENDACT 1030  
Gapop 10.0 , Gapext 0.5  
Scoring table: BLOSUM62  
Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5450	100.0	1030	1	SM6A_HUMAN
2	4346.5	79.8	888	1	SM6A_MOUSE
3	2013.5	36.9	988	1	SM6B_HUMAN
4	1937	36.6	657	1	SM6B_RAT
5	1935	36.6	886	1	SM6B_MOUSE
6	1523.5	27.8	930	1	SM6C_HUMAN
7	1493.5	27.4	931	1	SM6C_MOUSE
8	1431	27.4	960	1	SM6C_RAT
9	971.5	17.8	730	1	SM1A_SCHAM
10	932.5	17.1	771	1	SM1A_PRCME
11	859	15.7	772	1	SM3A_RAT
12	853	15.7	772	1	SM3A_MOUSE
13	652.5	15.6	771	1	SM3A_HUMAN
14	845	15.5	712	1	SM1A_TRICF
15	845	15.5	778	1	S21B_BRARE
16	845	15.5	860	1	S21A_BRARE
17	839.5	15.4	772	1	SM3A_CHICK
18	830	15.2	1374	1	SM3A_HUMAN
19	826	15.2	749	1	SM3B_HUMAN
20	825.5	15.1	1377	1	SM3A_MOUSE
21	803	14.7	748	1	SM3B_MOUSE
22	802	14.7	764	1	SM22_BRARE
23	797	14.6	777	1	SM3D_HUMAN
24	793.5	14.6	761	1	SM3C_CHICK
25	789	14.5	712	1	SM1A_CABEE
26	766	14.1	1053	1	SM5B_MOUSE
27	762.5	14.0	706	1	S22A_DROXHE
28	762	14.0	775	1	S23E_HUMAN
29	752	13.8	751	1	SM3C_CHICK
30	751	13.8	861	1	SM3C_MOUSE
31	750.5	13.8	697	1	SM2A_SCHGR
32	749	13.7	785	1	Q13275_homo sapien
33	744.5	13.7	785	1	Q88632_mus musculus

34	739	13.6	775	1	SM3E_MOUSE
35	737	13.5	751	1	SM3C_MOUSE
36	735.5	13.5	862	1	SM4D_HUMAN
37	733.5	13.5	785	1	SM3E_CHICK
38	733	13.4	751	1	SM3C_HUMAN
39	732.5	12.9	766	1	SM27_BRARE
40	692	12.7	834	1	SM4C_MOUSE
41	681	12.5	832	1	SM4B_HUMAN
42	672	12.3	761	1	SM4A_HUMAN
43	664	12.2	776	1	SM4F_RAT
44	658	12.1	777	1	SM4F_MOUSE
45	656	12.0	782	1	SM4B_MOUSE

ALIGNMENTS

RESULT :  
SM6A\_HUMAN  
ID SM6A\_HUMAN STANDARD; PR7: 1030 AA.  
AC Q9H2E6; Q9P2H9;  
DT 15-SEP-2003 (Ref.: 42, Created)  
DT 15-SEP-2003 (Ref.: 42, Last sequence update)  
DT 15-SEP-2003 (Ref.: 42, Last annotation update)  
DE Semaphorin 6A precursor (Semaphorin VIA) (Sema VIA) (Semaphorin 6A-1)  
DE (SMA6A-1)  
GS SEMA6A CR K2AAL368.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
[2]  
SEQUENCE FROM N.A., AND INTERACTION WITH EVL.  
RA MEDLINE=20564339; PubMed=1093894;  
RA Klostermann A., Lutz B., Gertler F., Behl C.;  
RT "The orthologous human and murine semaphorin 6A-1 proteins  
(SMA6A-1/Sema6A-1) bind to the enabled/vasodilator-stimulated  
phosphoprotein-like protein (EVL) via a novel carboxyl-terminal  
tyrosine-like domain."  
RT C. Biol. Chem. 275:39647-39653(2000).  
RN [2]  
PP SEQUENCE FROM N.A.  
PT TISSUE=Brain;  
RX MEDLINE=2019126; PubMed=10718199;  
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Chaza G.;  
RT "Prediction of the coding sequences of unidentified human genes. XVI.  
The complete sequences of 150 new cDNA clones from brain which code  
for large proteins in vitro."  
RT DNA Res. 7:65-73(2000).  
CC -1- FUNCTION: Can act as repulsive axon guidance cues. May play a role  
in channeling sympathetic axons into the sympathetic chains and  
controlling the temporal sequence of sympathetic target  
innervation (By similarity).  
CC -1- SUBUNIT: Active as a homodimer or oligomer. Interacts with EVL.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- ALTERNATIVE PRODUCTS:  
Event=Alternative splicing; Named isoforms=2;  
Name=1;  
IsoId=Q9H2E6-1; Sequence=D:played;  
Name=2;  
IsoId=Q9H2E6-2; Sequence=VSP\_007113;  
Note=No experimental confirmation available;  
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
CC -1- SIMILARITY: Contains 1 Sema domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



DEVELOPMENT.  
 C -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
 C -1- SIMILARITY: Contains 1 Sema domain.  
 C  
 C This SWISS-PROT entry is copyright. It is produced through a collaboration  
 C between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 C the European Bioinformatics Institute. There are no restrictions on its  
 C use by non-profit institutions as long as its content is in no way  
 C modified and this statement is not removed. Usage by and for commercial  
 C entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 C or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 C  
 R EMBL: AF030430; ARS6408.1;  
 R MGD: MG1123727; Sema6a.  
 R InterPro: IPR003652; Plexin-like  
 R InterPro: IPR003652; Sema.  
 R Pfam: PF01403; Sema; 1.  
 R SMART: SM00423; PSI: 1.  
 R SWART: SMC0630; Sema; 1.  
 R Signal: transmembrane; Multigene family; Neurogenesis; Glycoprotein;  
 W Developmental protein.  
 T SIGNAL 1 15  
 T CHAIN 19 888  
 T DOMAIN 19 649  
 T TRANSMEM 650 670  
 T DOMAIN 671 988  
 T DOMAIN 56 491  
 T DOMAIN 792 819  
 T CARBOHYD 33 33  
 T CARBOHYD 49 49  
 T CARBOHYD 65 65  
 T CARBOHYD 282 282  
 T CARBOHYD 434 434  
 T CARBOHYD 461 461  
 T CARBOHYD 899 AA; 96075 MW; C7094FC2EB345C57 CRO04;  
 Q  
 Query Match 79.8%; Score 4346.5; DB 1; Length 988;  
 Best Local Similarity 94.4%; Pred. No. 3-5e-272;  
 Matches 916; Conservative 24; Mismatches 23; Indels 1; Gaps 1;  
 Y 1 MRSEALLYTLHIFAGAPPEDEPISIGHGNTKCYPVVGHKGRNTTQRR:DIOM 60  
 b 1 MRPAALMLLCLLACAGAPPEDEPISIGHGNTKCYPVVGHKGRNTTQRR:DIOM 60  
 Y 61 TNINGTLYNARDHTVTVDTGSHBEIYCSKLTNKEBOAYGCRMKGKHORCHE 120  
 b 61 LGMNRTLYVAARDHTVTVDTGSHBEIYCSKLTNKEBOAYGCRMKGKHORCHE 120  
 Y 121 IKVLKKNDDALFVCGTNATPSCRNKYNKTTLEPPGDEFSGMACGYDAKHANVAFAG 140  
 b 121 KVLKKNKTLFVCGTNATPSCRNKYNKTTLEPPGDEFSGMACGYDAKHANVAFAG 140  
 Y 191 KYSATVTDFAIDAVTVRSLSGSEPTRTYKHSKMWKEEYFYQAVDYDYIYFFFEFA 240  
 b 181 KYSATVTDFAIDAVTVRSLSGSEPTRTYKHSKMWKEEYFYQAVDYDYIYFFFEFA 240  
 Y 241 VEYTMGVVFPVAVQCKNDYGSQRYLQKMTSELKARLNCVPGDSHFENLQAVT 300  
 b 241 VEYTMGVVFPVAVQCKNDYGSQRYLQKMTSELKARLNCVPGDSHFENLQAVT 300  
 Y 301 DVIRINGRDVJATFTSTPYNIPGSAVCAYMDLDIASVFTGRFKEQKSPSTWTFPDER 360  
 b 301 DVIRINGRDVJATFTSTPYNIPGSAVCAYMDLDIASVFTGRFKEQKSPSTWTFPDER 360  
 Y 361 VPKRPCCAGSSSLERYATSNRPDITLNP:KTHPLNDVAVSIFNRPWFLRTMYRYLJ 420  
 b 361 VPKRPCCAGSSSLERYATSNRPDITLNP:KTHPLNDVAVSIFNRPWFLRTMYRYLJ 420  
 Y 421 TKIADVDAAGPYQNTVYVGLSGRGGLILKE:ARTGNSGFLNGLSFLFEMSVYVSEKSYD 480  
 b 421 TKIADVDAAGPYQNTVYVGLSGRGGLILKE:ARTGNSGFLNGLSFLFEMSVYVSEKSYD 480  
 Y 461 GVEDKRMQMDRASSLYVAFSTCVIKVPLGRRCERHKGKKTCT-ASRDPYCGW-KEGG 540

Db 481 GVEDKRMQMDRASSLYVAFSTCVIKVPLGRRCERHKGKKTCT-ASRDPYCGWVRES 540  
 Qy 541 ACSHLSPNSRLTTEQDIERGNTDGLGCHNSFVALNGHSSLSLPSTTTSDS:TAQGYEGR 600  
 Db 541 SCAHLSPLSLTTEQDIERGNTDGLGCHNSFVALNGHASSLVPSTTTSDS:ASRQGYEGR 600  
 Qy 601 GGLMDKHLDDSPDTPFLGAVSSHHNQKKGKVIRESYKGGHDLQVPTVLLAIV:LAIV 660  
 Db 601 GGLMDKHLDDSPDTPFLGAVSSHHNQKKGKVIRESYKGGHDLQVPTVLLAIV:LAIV 660  
 Qy 661 YGAVESGITVYVCDHRRKQVAVVQSKKELTHSRGSMSSVTKLSGLGPTGCKSPKPE 720  
 Db 661 YGAVESGITVYVCDHRRKQVAVVQSKKELTHSRGSMSSVTKLSGLGPTGCKSPKPE 720  
 Qy 721 ALITPLMHNGKLATPNTAKMLIKAPQHLDLTALPTPESTTTLOKXKPSRGSEWERN 780  
 Db 721 ALITPLMHNGKLATPNTAKMLIKAPQHLDLTALPTPESTTTLOKXKPSRGSEWERN 780  
 Qy 781 QNINACTKDMPPGSPVPTDLP:PLBASPSHIPSVVVLPIQGGYCHYVDCPKMSE-VA 839  
 Db 781 QNINACTKDMPPGSPVPTDLP:PLBASPSHIPSVVVLPIQGGYCHYVDCPKMSE-VA 840  
 Qy 840 QMALEDQATLEKTKIKHEHLSKS 863  
 Db 840 QMALEDQATLEKTKIKHEHLSKS 864  
 RESULT 3  
 SM6B HUMAN STANDARD; PRI: 888 AA.  
 ID 09H3T3; C9NRK9;  
 DT 16-OCT-2001; Rel. 42; Created;  
 ST 16-OCT-2001; Rel. 40; Last sequence update;  
 ST 15-SEP-2001; Rel. 42; Last annotation update;  
 DE Semaphorin 6B precursor (Semaphorin Z) (Sema Z).  
 GN SEMA6B OR SEMAZ.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 PC TISSUE=Brain.  
 RA Kimura T, Ichida H.  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=2124868; PubMed=11350127;  
 RA Correa R.G., Sasahara R.Y., Bengtson M.H., Karayama M.L.H.,  
 RA Salim A.C.M., Brentani V.M., Sotayay M.C., de Souza S.J.,  
 RA Simpson A.J.G.  
 RA "Human semaphorin 6b".  
 RL Genomics 73:343-348(2001).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN BOTH PERIPHERAL AND CENTRAL NERVOUS  
 CC SYSTEM DEVELOPMENT (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=1;  
 CC IsoId=Q9H3T3-1; Sequence=Displayed;  
 CC Name=2; Synonym=6B.1;  
 CC IsoId=Q9H3T3-2; Sequence=VSP\_006044, VSP\_006045;  
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
 CC -1- SIMILARITY: Contains 1 Sema domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC



Query Match: 36.6%; Score 1997; DB 1; Length 887;  
 Best Local Similarity 42.7%; Pred. No. 7, 1e-121;  
 Matches 433; Conservative 132; Mismatches 263; Indels 193; Gaps 25;

1 INTERPTG; IPR003659; PLEXIN-LIKE.  
 2 INTERPTG; IPR001627; SEMA.  
 3 P1AM; PFC1403; SEMA; 1.  
 4 SMART; SM00423; PSI; 1.  
 5 SMART; SM00630; SEMA; 1.  
 6 Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;  
 7 Developmental protein.  
 8 SIGNAL; 26  
 9 CHAIN; 27  
 10 DOMAIN; 27  
 11 TRANSMEM; 606  
 12 DOMAIN; 627  
 13 DOMAIN; 597  
 14 DOMAIN; 549  
 15 SEMA  
 16 N-LINKED (GLCNAC); 75  
 17 CARBOHYD; 156  
 18 CARBOHYD; 168  
 19 CARBOHYD; 232  
 20 CARBOHYD; 387  
 21 CARBOHYD; 442  
 22 CARBOHYD; 463  
 23 SEQUENCE; 887 AA; 95752 MW; 05543P3F232CD3C1 CRC64;

2 REALLLYTLLHAGAGPSESEPISSHGYTKQYFVGHKKGENTTOR--HRLDIQ 59  
 10 RPALLFLLLLARVTHGSEDEFFPLSAPVDYLSHYFVGVSGGGRUTPAEGADLNQ 69  
 60 MIMVNGTLYIAARHIYVVDIDHTEBIYCSKLTWKSROADYDTPMKKQKDECHN 119  
 70 RVLRVNRTPFICERDNYGVLEPSTELKCRKLTPRNSFDIDVCRMKQKQGERH 129  
 120 FIKVLKNDADAFVCGTKAFPSQRYNMOTLRFQDFEFSQVACRYDAKNAVAFAD 179  
 130 FVKVLLPRESTLFVCGSNAPFICAKYSMTQLQLGKISQWACRYDPEKNAVAFSD 189  
 180 GLYSATVTDLAIDAVIVESGESPTLVKHSKMLKEPYFNGAVGVGVYFFPREI 239  
 190 GMLFTATVTDLAIDAVIVESGSRPTLVKHSKMLKEPYFNGAVGVGVYFFPREI 249  
 240 AVEYNWGVVPRVACVNDKMGQGVLEKQWTFKARLNCVSPVDSHYFVILCAV 299  
 250 ANEPVLEKVVSRVAVGVNDVCSGPRVLEKQWTFKARLNCVSPVDSHYFVILCAV 309  
 300 TVIRINGGVNLTSTPYNIFISAVCAVDYDLIASVFTGRFEQKQSPSTVTFVDE 359  
 310 TGWVSLGSRPVILAVFSTPSNIFISAVCAFDVQVAAVFEGRFEQKQSPSTVTFVDE 369  
 360 RYKPRPGGCASSSLERATSNPEFDTNEIKTHLPMDQAVPSIPRPFELRTVYR 419  
 370 QVPRPFGCAAPGM--QYNASVAFDEILAFVKTFLPMDQAVPSLGHSPWAVRTLRQ 427  
 420 LTKIADVTAAQPVNHTVFLGSEKILKFLAR--IGNSGELNDSLFEEVSVNKKC 477  
 428 LRVAVDVAGPVGNTVFLGSEVGVKFLKVPKMASVSGTGSILSEETVTPRSC 487  
 478 ---SYDGVTEKRMVQGLRASSLSLVAFSTVIVKVLQRCERHKKKTCASRPFYCG 534  
 488 GRSSSAGEGQRLLSLELAASGGLLAAPFCVAVFVAVACGLYSGCMKNC--GSQCPVCG 547  
 535 NIKEGGACSHLSPKSLTPEODIERNTGAGDCHNSFVALKCHSSLLPS--TTTDSIAQ 594  
 548 WAPD--GSCIFPREGTSATPEQVSGASTSGDGC----- 580  
 595 EGVESRGMGLDWKHLLDSPDSTDFLQAVSSHNQDKKGVIRESYLKGHDQLVPVTLAIA 654  
 581 -----TGJJRASLSDDRAGLVSNLLVTS 604  
 655 VILAFVNGAVFGSGITVYCVCDHRRKQDAVAVRKEKE--LTHSRGSMSSVTKL----- 705  
 605 SVAAFVVGAVSGVSGVGVGLRERRELA--RRKQKEALAHGGSEAVLSVSRJGERRGT 662

CY 766 --SGLFGTQSKQPKBEALITPLMHNGKLTATPNTAKMLKADCHHDLTALTPPSTPT 763  
 DB 663 GTGRCGAGGGGPGFPEALAPLMQNGV-----TKAALLHGGPHSDSGLLPTPQTP- 715  
 CY 764 LOOKRXFSPR-----GSREMERNQNLINACTK-----CXPNGSGPVIFTD-- 802  
 DB 716 LPQKRUPPTPHAHALGPPRANOHSHALLSASASTSLLLAHTPEQPP-----VFTEG 770  
 CY 903 -----LPLRASPSPHPSVVVLDITTCGYQHYEVDPKMSVAQVALEDOAATLEYKTIK 856  
 DB 771 PESRLCAPRSCRASHPGDFELTP-----HASPDRRRVVSAPTGPLDSSSVG----- 816  
 CY 857 EHSSSKSPNHGVNLENLDSLPPKVPQCEASL-----CPPCASLCTGLSKKLEMHHS 909  
 DB 817 -----DCLPGWSPPTATSSURRPGHPPTAAARRT-----H 845  
 CY 910 SSVGVGVYKRSYPTKSLTFSHQATTAKRNTNNSNSSLRNQSGFRGDDNPAPP 963  
 DB 950 FNSG----EAPGGRHPRHA-----PADSTHL---LPCGIGERTAPPVP 957

RESULT 5  
 SM6B MOUSE  
 ID SM6B MOUSE STANDARD; PRT; 886 AA.  
 AC 064951;  
 DT 30-MAY-2000 (Rel. 33, Created;  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 43, Last annotation update)  
 DE Semaphorin 6B precursor (semaphorin V-B) (Sema VIB) (Semaphorin N)  
 DE (Sema N)  
 GN SEVA6B CR SEMAN  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 FN NCBI\_TaxID=10090;  
 RN [1]  
 PP SEQUENCE FROM N.A.  
 EX MEDLINE=98027184; PubMed=9361278;  
 RA Eckhardt F., Behar O., Calautti E., Yonezawa K., Nishimoto I.,  
 RA Fisman M.C.;  
 RT "A novel transmembrane semaphorin can bind c-src.";  
 RL Mol. Cell. Neurosci. 9:459-419 (1997);  
 CC -- SUBUNIT: HOMODIMER. BINDS SPECIFICALLY THE SH3 DOMAIN OF THE  
 CC PROTOONCOGENE C-SRC.  
 CC -- SUBCELLULAR LOCATION: Type : membrane protein.  
 CC -- TISSUE SPECIFICITY: DURING DEVELOPMENT IT IS EXPRESSED IN  
 CC SUBREGIONS OF THE NERVOUS SYSTEM AND IS PARTICULARLY PROMINENT IN  
 CC MUSCLE. IN ADULTHOOD, IT IS EXPRESSED UBICUITOUSLY.  
 CC -- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
 CC -- SIMILARITY: CONTAINS 1 Sema domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EVBL; AF366855; AAC00493.1; --  
 DR MGI; MGI:1202889; Sema6b.  
 DR InterPro; IPR001659; Plexin-like.  
 DR InterPro; IPR001627; Sema.  
 DR Pfam; PFC1403; Sema; 1.  
 DR SMART; SM00423; PSI; 1.  
 DR SMART; SM00630; SEMA; 1.  
 KW Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;  
 KW Developmental protein.  
 FT SIGNAL 26  
 FT CHAIN 27 886  
 FT DOMAIN 27 605  
 FT TRANSMEM 606 626

T DOMAIN 627 986 CYTOLASMIC (POTENTIAL).  
 T DOVAIN 239 549 SEMA.  
 T DOVAIN 751 754 POLY-JEU.  
 T CARBOHYD 75 75 N-LINKED (GLCNAC) (POTENTIAL).  
 T CARBOHYD 156 156 N-LINKED (GLCNAC) (POTENTIAL).  
 T CARBOHYD 292 292 N-LINKED (GLCNAC) (POTENTIAL).  
 T CARBOHYD 387 387 N-LINKED (GLCNAC) (POTENTIAL).  
 T CARBOHYD 442 442 N-LINKED (GLCNAC) (POTENTIAL).  
 T CARBOHYD 463 463 N-LINKED (GLCNAC) (POTENTIAL).  
 Q SEQUENCE 886 AA; 95466 MW; 58560125CDAS74D CRC64;

Query Match 36.63; Score 1995; Da 11; Length 886;  
 Best Local Similarity 43.74; Pred.No. 9.5e-421;  
 Matches 438; Conservative 132; Mismatch 257; Indels 176; Gaps 28;

Y 8 LVFLLHFAAGAG---FPEDSFIIISHQNTKQVYFVGHKPGKNTQCR--HELDQMY 62  
 b 13 LSPFLLLLGTYGLFPEEPPLVAVRQVLSHYFVFGSPRLTAAAGADQNIQVYL 72  
 Y 63 IMNGTVAARDHVTYDITSHTEIYCSKKLTKWKSQAFVDTCKMKKHQDECHNPK 122  
 b 73 RWRKTLFGDENLVQVLEPSTGTELVYCRQVWAGNTSGIDVCKMKKQSGCKNFKV 132  
 Y 123 VLLKNDQALFVQGTNAFNSCRNYKNTLEPQDQPSQWACVYAKUNVAFAGKYL 182  
 b 133 VLLRDESTLFVQGSNAFNPICAYSVETLQAGGIGMARGTYDQKIANVAIFEDGKL 192  
 Y 193 YSATVTFPLADAVYPSLDESPTLRVWQSKWKLKEPFVQAVQVGYVFFFEIAYE 242  
 b 193 FTATVTFPLADAVYPSLDEPTLRVWQSKWKLKEPFVQAVQVGYVFFFEIAYE 252  
 Y 243 YNTMKVYFVAVQVCKNDGSGORVLEKQWTFELKARLNCVPGSHFYFNLQVTVY 302  
 b 253 FNYLEKVVYFVAVQVCKNVGGSHRVLEKQWTFELKARLNCVPGSHFYFNLQVTVY 312  
 Y 303 IRIKGRDVLATPSTPNYSIPGSVAVCAVMDLTASVFTGRPKCKSPDSTWTPVDERVP 362  
 b 313 VSLGGRPVILAVPTPNSIPGSVAVCAVMDLTASVFTGRPKCKSPDSTWTPVDERVP 372  
 Y 363 KPRGCGAGSSSLERYATSEFFPTLNFVKTPLMLDEAVPSIPNPFVLMVRYLTK 422  
 b 373 RPRGCGAARPYQYNASSALPEIENFKTHPMDXDEAVESLGHSEFWITLMIHQLR 430  
 Y 423 IAVTAAQVQVNHVYFVSGKSGVLLKFLAR--IGKSGFVNDLSLFEEMSVNSEKC--- 477  
 b 431 VAVDVGAGVWENQVIFLGEAGVLLFVFKPKNASVSGTTPGSLFEEFETVAPDQGRF 490  
 Y 478 SYGVGVEDKRIKNGCLDRASSLYVAFSTGVKVPVGRCEBHGKCKXTCTIASRDPYGMVK 537  
 b 491 SSAGWGQRLSLBELDAASGGLLAAPPVQVVPVAVACQVSGCKNKGISQDPYSGWAP 550  
 Y 538 EGGACSHLSKSRUTERQF--ERQNTDGLGRCHNSFYALNCHSSSLSPSTTSQAEY 597  
 b 551 D-GSCFLRPPSATFQDVGAGTSGSGDC----- 580  
 Y 598 ESRGMLDWKHLSDPSSTDPVGVAVSHNHQCKGVIRESVLXGHQDLVPTLLAIVL 657  
 b 591 -----TGULPASLSEDFACGLSVNVLVTSVA 607  
 Y 659 AFVMAVPSGITVYCYCDHRKCVAVVQKKEK--DTHSRGSMGCVKLSCLPFCDS 714  
 b 658 AFVGVAVVSGVGVGVFVQVRRERELA--RRCKPALLAHAGSEAVLSVSRV---GERS 761  
 Y 715 KDP-----KPAILTPIIMHGKLTATGNTAKVLIKACQCHDLTAFTPESTP 762  
 b 662 TGPGRGAGGAGGPGFPALLAPINQW-----TKAAILHGPGIDLTGSLTPTPECP 715  
 Y 763 TLOQKKRP-----SRGSEWENONLNACTKQNPFGSPVPIPTL-----PLRASPGH 811  
 b 716 LPQKELTPPHRAHAGLRANDSHALLSAS-----NSLLAPARASEQ- 762  
 Y 812 IPSVVVLPITQGGYCHEVDQF---KQSEVACMALEDQALLEVYIKHLSKSKSNHGV 869  
 b -----

Db 763 -POVPAEPGPE-----SRCAPSCASCRASHPDPFLTPHASPDPRRWVSAPTGLDPSVG- 815  
 Qy 869 NJVENLDLP-PKYPOFEASL-----GPGGASLCTGSKLEHSHSSYGVYKRSYP 921  
 Db 916 -----DGLPGASPPATSSLRPPGHPPTAALRR-----HTFNSG-----EARP 856  
 Qy 922 TNSLTSHQCATLKRKNTSKSSSLSRNQSPGRG3-NPEEPAP 963  
 Db 857 GSHRPRHP-----PADSTHL---LPGTGERTAPPVP 886

RESULT 6  
 SM6C\_HUMAN  
 ID SM6C\_HUMAN STANDARD; PRT: 93C AA.  
 AC Q9H3T2; Q9H3T2; Q9H3T2; Q9H3T2;  
 DT 16-OCT-2003 (Ref. 40; Created)  
 DT 28-FEB-2003 (Ref. 41; Last sequence update)  
 DT 28-FEB-2003 (Ref. 41; Last annotation update)  
 DE Semaphorin 6D precursor (Semaphorin Y) (Sema Y).  
 GN SEMA6C OR SEMAY.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RK [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC [1]SSEA-Brain;  
 RA Kimura T., Ishida H.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
 RC [1]SSEA-Brain;  
 RA Cai X., Zhai Y., Wei H., Zhang C., Xing G., Yu Y., Wu S., Zhang Y.,  
 CC Yang S., Zhou G., He F.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 CC [1] FUNCTION: MAY BE A STOP SIGNAL FOR THE DORSAL ROOT GANGLION  
 CC NEURONS IN THEIR TARGET AREAS, AND POSSIBLY ALSO FOR OTHER  
 CC NEURONS. MAY ALSO BE INVOLVED IN THE MAINTENANCE AND REMODELING OF  
 CC NEURONAL CONNECTIONS (BY SIMILARITY).  
 CC [1] SUBCELLULAR LOCATION: Type I membrane protein.  
 CC [1] ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=1; Synonyms=Short 1;  
 CC Name=2; Synonyms=Short 2;  
 CC Name=3; Synonyms=Long;  
 CC [1] SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
 CC [1] SIMILARITY: Contains 1 Sema domain.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AB022434; BAB20870.1;  
 CC EMBL; AF339152; AAL72039.1;  
 CC EMBL; AF339153; AAL72039.1;  
 CC EMBL; AF339154; AAL7200.1;  
 CC GenBank; HGNC:10740; SEMA6C.  
 CC InterPro; IPR003659; Sema-like.  
 CC InterPro; IPR001627; Sema.  
 CC Pfam; PF01403; Sema; 1.  
 CC SMART; SM00423; PS; 1.  
 CC SMART; SM00430; Sema; 1.  
 CC Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;  
 CC Developmental protein; Alternative splicing.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 930 SEMAPHORIN 6C.

DOXAIN 25 604 EXTRACELLULAR (POTENTIAL).  
 TRANSMEM 605 525 POTENTIAL.  
 DOXAIN 626 230 CYTOPLASMIC (POTENTIAL).  
 DOXAIN 233 343 SEMA.  
 DOXAIN 662 667 POLY-PRO.  
 DOXAIN 752 755 POLY-PRO.  
 CARBOHYD 70 70 N-LINKED (GLNAC).  
 CARBOHYD 286 286 N-LINKED (GLNAC).  
 CARBOHYD 437 437 N-LINKED (GLNAC).  
 VARSPLIC 184 223 Missing (in isoform 2).  
 VARSPLIC 566 566 /FTID=VSP\_066046.  
 Y -> YVJGPGSPGTPSPDPAHFRPQSSTLGHTR  
 (in isoform 2 and isoform 3).  
 /FTID=VSP\_006047.  
 I -> V (IN REF. 2; AAL70299).  
 R -> K (IN REF. 1).  
 P -> T (IN REF. 1).  
 SEQUENCE 930 AA; 99652 YW; 8AF8814ADBC84C98 CRC64;  
 Query Match 27.8%; Score 1515.5; DB 1; Length 930;  
 Best Local Similarity 35.8%; Pred. No. 7.7e-90;  
 Matches 380; Conservative 136; Mismatches 356; Indels 189; Gaps 32;  
 6 LLYFTLHFAAGAPEDSESPIS-SHKNYTKYCPVFGKFKFGRTTCRRRLCIQXIMN 65  
 13 LLLLLSLPH-TQAAFPQDPLLLSLQGTSLFSLWFRGLEDDAAVEL-GDFQRFELTN 70  
 66 GLVY-AARCHYTYDIDT-SHT-BEYCSKKYTWKBRQADVDTCMKKGKKECHNFKNV 124  
 71 RLLVNAHDHVFSPDQAEBEGLVPNKYKFWESQ--DVNCAVRGKLIDECVYIRLV 129  
 125 LKNDLALFVGTNAFNPSCNRYKVTLEPPEDFSPGVARCFYDAKHANVALFADGKLYS 184  
 129 VPWDSQTLACGTSFSPVCSYSGYTSIQQGEELSGQARCFPATGCVAFLEGSLYS 188  
 135 ATVYDFLAIDAVIVASLSESPYLPYTWKSDSKMLKEPYFQVQVQVGYIYFFERLAVEM 244  
 189 ATAADFQASQAVVVASLQFPPLRSKAYDSKMLKEPYFQVQVQVGYIYFFERLAVEM 248  
 245 TWGKYTFYPRVACVCHDGGGQORVLEKONTFLKARLNCSPVDSHFENLQAVTVIR 304  
 249 PLGRVCFGRVAVCKEDYGGSPRAEDRWTSFLKRLNCRVGGSTFYVFLQALSTQPN 308  
 305 INGVQWLATSTPYNS-PGSAVCAYMCLDASVFGFKCKECSFDSWTVPVDPVPKP 364  
 309 LHGRNALFGVTTQNS-PGSAVCAYFLDELRGFKGKESRDLGAWTVSDRPPSP 369  
 365 RFGCAGSSSLERYATSNFEDTLNFKTHPLMPEAYPSFENRPFVLTWVRYELTKIA 424  
 369 RFGSCAGVGAALFSSRDLFDVULTFKAHFLDFAVPPVTHOP-LITLSSRALLTQVA 427  
 425 VDTAGPYQNTWVFLGSEKGLIKFLARLNGSPFNDSPLEBNSVYNSKCKYGVGD 484  
 426 VGVAGPHSNITVWFLGSDGTVLKVLPFGRRSG-PEPILLESIDAVSPARCSKTAQ 486  
 485 --KPMGMHQLPRASSSVAVSTCVIKPLGRCHGKCKTKCLASRQPYCGWIKERGA 542  
 487 TARR--GLIEDTEGRLVFAVSSGCVYPLPSCRARHGACQPSCLASQPYCWNSSRG 545  
 543 SHLSNSRLTRFQIERGNTDQ--GDCHNSFVALNGHSSLLRSTTSDTAQFGYER 600  
 546 VDIRSGGTVDQ--AGNQSMERHGDQDQ--ATGQSGPGPSAY--- 586  
 601 GGLMDKWKLLDPSDTPD-LGAVSNHQQKGVITRESYLKHKDQVFTVLLAIVLAFV 660  
 587 -----GVRRLPPASASRSVPILPLLASVAAPA 615  
 661 MGAVFSGITVYCVCDHRKQVAVVCRKEKELTHSPRG-----SKSSVTKLSGL 708  
 616 LGASVSGLLVSCAC--RR-----AHRRRGKOITPSPRLPLSRSJARLHG- 659  
 709 FGDTSKQPKP--EAILPLMHNGKLATPGNTAKMLKADCHHDLTALP-PESTFTLQJ 766

660 -GGPEPPEPESKQDQAVQITGLYTFLLPPEEGVPP-----ELACLPTESTELPV 709  
 767 KRPDSRGSSEKERNONLINACTKMPNMSFVPTDPLRASPSH-----IPSVVLPIT 821  
 710 KHLRAAGD-PWENWONRNA-----KEGFGSRSGHGAAGGAPAPRVLYRP-- 752  
 822 QCGYQHEYYDQPMKSEVACMALEDQAATLEYKTIKHL-----SSKSPNKGVLNVLDS- 876  
 753 -----PPGCGFGC-----AVEVTLELLYLHGPQPPRKGAEPAPLTSR 793  
 877 -LPKVPQSEASLGPFGASLSQTLGSKRLEMHSSSYGVYKESYPTNSUTRSHQATLK 935  
 794 ALPPE--PAPALGGSPRPHECASPRLDV-----PPEGRCASAPA---- 833  
 936 RNNTNSNSHL-----SRNQSFGRGNGPPAPORVDSIQVHSSQFSGQAVTVSRQPS- 988  
 834 -RPALGAPAPRLGVGGRALPFGSHRAPFALLTRV-----PSGSPSRYSGGPGKHL 883  
 989 --LNAYNSLTRSGLKRTPLSKPDVPPKPSF-APLSTSMKPN 1026  
 884 -YLGRPEGVGRALKRVDVEKPOLSLKPLVLGPPSSRCQAVPN 924  
 RESULT 7  
 SM6C VCUSE STANDARD; FRT; 931 AA.  
 ID SM6C VCUSE  
 AC QWTV2  
 DE 30-MAY-2000 (Rel. 39, Created)  
 DE 30-MAY-2000 (Rel. 39, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Semaphorin 6C precursor (Semaphorin Y) (Sema Y).  
 GN SEMA6C OR SEMAY.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN NCBI\_TaxID=10090;  
 RE SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=9916821; PubMed=1049528;  
 RA Kikuchi K., Chedotal A., Hanafusa H., Ujima Y., de Castro F.,  
 PA Goodman C.S., Kimura T.;  
 PT Cloning and characterization of a novel class VI semaphorin,  
 PT semaphorin Y.";  
 PL Mol. Cell. Neurosci. 13:9-23(1999).  
 CC -- FUNCTION: MAY BE A STOP SIGNAL FOR THE DORSAL ROOT GANGLION  
 CC NEURONS IN THEIR TARGET AREAS, AND POSSIBLY ALSO FOR OTHER  
 CC NEURONS. MAY ALSO BE INVOLVED IN THE MAINTENANCE AND REMODELING OF  
 CC NEURONAL CONNECTIONS (BY SIMILARITY).  
 CC -- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
 CC -- SIMILARITY: Contains 1 Sema domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: A5013729; BAA74294.1;  
 DR MGD: MG21338032; Sema6c.  
 DR InterPro: IPR001627; Sema.  
 DR Pfam: PF01483; Sema; 1.  
 DR SMART: SM00630; Sema; 1.  
 KW Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;  
 KW Developmental protein.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 931 SEMAPHORIN 6C.  
 FT DOMAIN 26 605 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 606 626 POTENTIAL.  
 FT DOMAIN 627 931 CYTOPLASMIC (POTENTIAL).



T DOXAIN 234 541 SEXA.  
T DOMAIN 663 669 POLY-PRC.  
T DOMAIN 754 757 POLY-PRC.  
T CARBOHYD 7 7 N-LINKED (GLCNAC) (POTENTIAL)  
T CARBOHYD 287 287 N-LINKED (GLCNAC) (POTENTIAL)  
T CARBOHYD 438 438 N-LINKED (GLCNAC) (POTENTIAL)  
Q SEQUENCE 931 AA: 93537 MW: 51099594205F125 CR264;  
Query Match 27.4%; Score 1493.5; CB 1; Length 931;  
Best Local Similarity 35.7%; Pred. No. 2e 88;  
Matches 374; Conservative 145; Mismatches 153; Indels 177; Gaps 32;  
Y 6 LLLVFTLLHFRAGAGFPEDSPIS:SHQNYTKQTVYVGHKFKRHTTORHRLDQIMKMM 65  
b 13 LLLLSLSPQAQAAPPQTPPLSLQLQASPSWFRGLEDDAVALSLGLSPQFQFLM 71  
Y 66 GTVYIAARDHYTVDTTSHT-EMIVSKSLTKWSQAGVDTCKMKGIKQCHNFKV 124  
b 72 RLLVAARDHYVFDLQACQCEGLVFNKELTRSQ--DMKCAVAGSKLTDGYNVIRV 129  
Y 125 LKXEDALFVGGTNAHPSQNYMMOTLHPGEPFGMACTYDAKIANVALFAQKLYS 154  
b 130 VPMKSTQLLACGTSNPSMCKRSYGLTSLQCGEELSQCAQCPFRATQSTVA:FAEGSLYS 189  
Y 155 ATVTDFALDAVIVRESIGESPTLTHKIDSKW:KEPYFOAKYCYGVYVFFREIAVENY 244  
b 190 ATAADPQASAVVHSLSPQFFLNSAKYDSKWRBHFVYALEHGHVYFFFRKSVRMA 249  
Y 245 TMGVVVFVFAVCKNGMSQRLEKQCTSLFKARLNFVFDSSHYPNLIQANTDVF 304  
b 250 RLGRVQFSVARVCKXNGSPALCBHW:SLFKLRINCVSQDSTFYEDVLQSLGPN 309  
Y 305 INGRVVAFTSPYNS:FGSVCAYDMLDIAVFTGRKEQASPSHTWIPDGRVAP 364  
b 310 LHGRSALEFVFTQTSNPSAVCAFYDCTERQFGKPKQCSLQCAVPSYEDKVPSP 369  
Y 365 RPGCAGSSSLERATSENPDDTLNFKTHPPMDKAVPSIKRPFRTMVPYRTKIA 424  
b 370 RPSGAGVGAASASSQDLPDVLFIKHPLLDPAVPPATHQP-LLTLSPALLTQVA 429  
Y 425 VDTAAGPYONHTVFLGSEKGIIFKFAIGNSGFLNLSLFJEMSVYNSKCSVDGVED 454  
b 429 VDGXAGPHRNTVFLGSDGCTVLKVLPP-GGQSLGSEPTVLEIDVAYSHRCS--GRKS 455  
Y 485 ----KXNGVCLDASSLYVAFSTCV:KVP:GRCEPHGKATCTIADSPVCGM:KEGS 542  
b 486 PRAARRIGLELDTGHRJLVAFFGCTVYVLSGRCAHGAQCSGLASQSPYCGMHRSG 545  
Y 541 ACSHLSFNSRLTFEQDIERGNTDGLGCHKSFVALAGHSSLLPSTTSDSTAQEGYE-S 599  
b 546 QMSIRGP-----GGTD-----VDLIGNQES-----TEHGDCQDQATGS 578  
Y 600 RSGVLDKXHLDSQDSTDP:GAVSSHNHQKGVIRBSYLGKHQDLVPVTLIAVILAF 659  
b 579 QSGFGDSAY-----GVRDLSFASASRSLP:PLSLACVAAAF 615  
Y 660 VYGAVFGSTVYVCDHRKDAVAVORKEKELTHSRGGSVSVTKLSGFGDTQSDKRP 719  
b 616 ALGASVGLLVSCAC--RRAN----RRSKDLETPG:FRPLSRSLARLHGSPSPFP 669  
Y 720 ---BALUTPLMHNGKLATPGNTAMJLNADQHLDTALTPTSTPTLQQRKPSGRSE 776  
b 670 KDGAAQTQEQVITLFPPOGASFP-----ELACLPTPTTTELFVXK:RAGSG-P 719  
Y 777 KERNQNLINACT-KDPMGSGFVPTDLP-----LEASFSHIPSVVVF:FIQQ--GYCH-- 827  
b 720 KEMNQGNHASEGFRPPGSGGAPRVLPVAPPPGCGQAVEVTTUEBLRYLRSP 779  
Y 828 -----EYVQPKMSVACMALEDQAATLEVKTIKELSSKGFPHGVNLVNL 874  
b 780 QPPRKGSPLASAFSTSPRPASEFASGLFVD-----SSPMPGCVPL-PL 824  
Y 875 DSLPFPK-----VFQREAS:GPPGASLSQTGLSKRLMHHSISYGVYKRSYPTNLSLTSHQ 910

D5 D-VPPEGTRAAPSGRFALGAPAPRLG-VGSGRL-----PFPT-----HR 862  
QY 931 ATTLKRNTNSNSHLSRNSFGRCNRPAPQVRVDS--IQVHSSQPSQAVTVSRQPS 963  
D5 863 A-----PFGLLTRVPSGPARYSGGPRHLLYLGR-PE 894  
QY 989 INAYNSUTRSGLKRTPSLKPQV--PPKPS 1015  
D5 895 GHRGRSLKRVDKVSP:SPKPLASPPQPA 923  
RESULT 5  
SM6C PAT STANDARD; FRT; 96C AA.  
AC Q9WTF3; Q9WTF6;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
GN Semaphorin 6C precursor (Semaphorin Y) (Sema Y).  
CS Sema6C OR SEMAY.  
CC Rattus norvegicus (Rat).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
CC NCB:TaxID=10115;  
RN 1;  
PP SEQUENCE FROM N.A. (ISCPORVS SEMA Y-L AND SEMA Y-S).  
RC STRAIN=Sprague-Dawley; TISSUE=Muscle;  
RX MEDLINE=9516082; PubMed=1049528;  
RA K:Kuch: K.; Chedtal: A.; Hanafusa R.; Ujinasa Y.; de Castro F.;  
RA Goodman C.S.; Kimura T.;  
RT "Cloning and characterization of a novel class VI semaphorin,  
semaphorin Y.";  
RC Cell. Neurosci. 13:9-23(1999).  
CC FUNCTION: SHOWS GROWTH CONE COLLAPSING ACTIVITY ON DORSAL ROOT  
GANGLION (DRG) NEURONS IN VITRO. MAY BE A STOP SIGNAL FOR THE DRG  
NEURONS IN THEIR TARGET AREAS, AND POSSIBLY ALSO FOR OTHER  
NEURONS. MAY ALSO BE INVOLVED IN THE MAINTENANCE AND REMODELING OF  
NEURONAL CONNECTIONS. Type I membrane protein.  
CC SUBCELLULAR LOCATION: Type I membrane protein.  
CC ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=Sema Y-L;  
CC IsoId=Q9WTF3-1; Sequence=Displayed;  
CC Name=Sema Y-S;  
CC IsoId=Q9WTF3-2; Sequence=VSP\_066048;  
CC TISSUE SPECIFICITY: EXPRESSED IN MANY REGIONS OF THE DEVELOPING  
NEVUS SYSTEM, PROBABLY IN NEURONS AND THEIR PRECURSORS, BUT ALSO  
IN NONNEURAL TISSUE SUCH AS IMMATURE MUSCLE AND DERMIS. IN ADULT,  
STRONG EXPRESSION IN THE SKELETAL MUSCLE AND MODERATE EXPRESSION  
IN THE BRAIN, WHERE CEREBELLUM SHOWS THE HIGHEST EXPRESSION. ALSO  
EXPRESSED IN ALMOST ALL AREAS OF THE CNS.  
CC DEVELOPMENTAL STAGE: DETECTED AT E1.2 AND FOUND AT MARKEDLY  
INCREASED LEVELS AT E1.5 AND E18 IN BOTH THE HEAD AND THE BODY. AT  
BIRTH THE LEVEL DECREASES SIGNIFICANTLY.  
CC SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
CC SIMILARITY: Contains 1 Sema domain.  
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL. Distribution  
of the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: AB003817; BAA76293.2; -;  
CC EMBL: AB014074; BAA76295.1; -;  
CC InterPro: IPR001627; Sema.  
CC Pfam: PF01403; Sema; 1.  
CC SMART: SM00630; Sema; 1.  
CC Signal: Transmembrane; Multisign family; Neurogenesis; Glycoprotein;  
CC Developmental protein; Alternative splicing.  
KW



```

T DOMAIN 215 515 SEXA.
T DOMAIN 217 222 POLY-PHE.
T CARBOHYD 44 44 N-LINKED (GLNAC. . .) (POTENTIAL)
T CARBOHYD 71 71 N-LINKED (GLNAC. . .) (POTENTIAL)
T CARBOHYD 163 163 N-LINKED (GLNAC. . .) (POTENTIAL)
T CARBOHYD 267 267 N-LINKED (GLNAC. . .) (POTENTIAL)
T CARBOHYD 360 360 N-LINKED (GLNAC. . .) (POTENTIAL)
T CARBOHYD 539 539 N-LINKED (GLNAC. . .) (POTENTIAL)
O SEQUENCE 730 AA: 8,2124 MW: 62625946576857F CPC64.

Query Match 17.83; Score 971.5; DE 1; Length 730;
Best Local Similarity 34.43; Pres. No. 5,9e-55;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

Y 8 LYFTLLHFAAGPDEPDEP-SISGNTVTKYPPYVGHKKGRNTQP-----RLDIQVI 61
b 11 LILWALH--AAAWNDISP-----KVVQF-----GEERVQDFLNGSKHCKFKLL 54
Y 62 KIMNGTUYLAARDHIYVD-D--TSHTETTYGSKKTKWSPQADVDYTKYGHKDECHN 119
b 55 EKDRNSLLVGARNVWISLRDUTETPEQ-----RTEWSSGARRRLCYLKGSESDYCN 109
Y 120 FIKVLLKKDGLPVCNAPKPSGNYKDTLEPQD ---EPGVARCPYDAGWY 174
b 110 YRVLAIEDRVDLCOTNAYKPLCRHYALKD-----GLYVWEYETGRGLCFEDPHNHT 164
Y 175 ALFAEGRLYATVDFATCAVIVSLSGESFTLRTYTHSKKKKEFYFQAVDYDHYTFP 234
b 166 AIYEGGLYATVADPGSTDLIVFG-----PLTERSDKQLNAPNVTWYETCFIFP 220
Y 235 FFRRIAYVYTMGVPRVACVYKNDGSSQVLEKCTSHFKARLNSVGDQSHFVN 294
b 221 FFRFAYEYVNOGKAIYVARVCKDKGHGFGFR-WTSHLKLNSGVGDQSHFVN 279
Y 295 ILOAVTDWIRN-GRDY---VLAFFSPYNSIGCSAVCAVCMUJLASVTGFRKESKD 350
b 280 EICSTSDIEGVGCVQVEKLYGVFTTPNINSIGSAVCAESMSILEDGFRKCEETKN 339
Y 351 STWTPVDERVPRKPCGCGSSLEAYATSNFPDPTLNFTKTHPLMEATPSIFKRW 410
b 340 SNMLAVPSLKVPRPRGQCVND-----SRLPDVSNVFNASHTLDEAVFAFTRPI 391
Y 411 FLRTMVPRITKIAND---TAGPYQNTVTLGSEKGLLAFELARIGNSFLN----- 461
b 392 LIRLSGHRFKIACQCVNTPQD--KAYDVLFGIDLGKVIKAL-----NSAFSSDTV 445
Y 462 DSLFLEKSYNNEKSGYGVDEHRLVGMOLDRASSSLVAFSTGVKVPJGRC--EREG 519
b 446 DSVVIELQVLP-----PGVPVNLVWMDGSDSKLVVSDDELAIKLHRCGSKIT 459
Y 520 KCKKTCASRDYPCGNIKESGACSL-SPN---SKUTFEQDIERGNTDGLGDCHNSFVAL 575
b 500 NCRE-CVSLQDPYCAWNLKCTAVGSPQWAKGRFFIONSLGHEKACGGPQCEIV- 557
Y 576 NGRSSSLPSTTSDSTA-----QGVESRSGXLDKWKHLSDSPSTPLGAVSSRHHQ 626
b 558 ---ASPVPTPTTKSCDPVSHIQHAEFPE---IDNEIVIGVDSNVIPRTLALINHA 610
Y 629 DKGVIRESYLKHGDQVVPV---TL-LAI-----AVLIAFYGVGVFS 666
b 611 GSK-----LPSSQEKLPITYAETLTATVTSIGALVGVFSGFIFS 652

RESULT 10
41A DROKE STANDARD; PRT; 77: AA.
T 30-MAY-2000 (Rel. 39, Created)
T 30-MAY-2000 (Rel. 39, Last sequence update)
T 16-OCT-2001 (Rel. 42, Last annotation update)
S Semaphorin 1A precursor (Semaphorin-1) (Sema 1).
S Sema-2A OR DSEMA-1 OR CG18405.
S Drosophila melanogaster (Fruit fly).

```

```

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
[1]
SEQUENCE FROM N.A.
MEDLINE=94094332; PubMed=6269517;
MEDLINE=94094332; PubMed=6269517;
"the semaphorin genes encode a family of transmembrane and secreted
growth cone guidance molecules.";
Cell 75:1389-1399(1993);
SEQUENCE FROM N.A.
MEDLINE=20196006; PubMed=10731132;
Adams M.D., Celisner S.E., Holt R.A., Evers C.A., Sccayne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman M.R., Yandell M.C., Zhang Q., Chen L.X.,
Brannon A.C., Rogers Y.-H.C., Blazer P.G., Champe M., Pfeiffer R.D.,
Wan X.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., McKios G.L.G.,
Abell C.P., Agbayani A., An H.-C., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.Y., Basu A., Bayovale J., Beckett L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman S.P., Bhandali D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
Burris K.C., Busa D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Fabros B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.X.,
Dodonson J., Douc E.E., Downes M., Dugan-Rocha S., Durkin B.C., Dunn P.,
Dustin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Folier C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
Gluck A., Gong F., Gorrell C.H., Gu Z., Guan P., Harris M.,
Harris K.L., Harvey D., Heiman T.C., Hernandez C.R., Hock C.,
Hoskins R.A., Houston K.A., Howland T.J., Wei M.-H., Ibegwar C.,
Jatali V., Kalush F., Karpen G.H., Ke Z., Kennison C.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko F., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Jin X., Mattet B., McIntosh T.C., McJeed M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mout R.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
Palazzo J., Pittman G.S., Pan S., Pollard C., Puri V., Reese M.G.,
Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shen B.C., Siden-Kiamos S., Simpson M., Skupski X.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Swarg K.R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
Williams S.V., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.C.;
The genome sequence of Drosophila melanogaster.;
Science 287:2185-2195(2000).
-!- FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE.
-!- SUBCELLULAR LOCATION: Type : membrane protein.
-!- TISSUE SPECIFICITY: EXPRESSED BY SUBSETS OF NEURONS AND MUSCLES.
-!- DEVELOPMENTAL STAGE: EXPRESSION BEGINS AROUND STAGE 19, PRIMARILY
IN THE DEVELOPING CNS. IN STAGE 16 EMBRYOS, IT IS EXPRESSED AT
HIGHEST LEVELS THROUGHOUT THE CNS, AND WEAK EXPRESSION IS SEEN IN
PORTIONS OF THE PERIPHERAL NERVOUS SYSTEM, MOST CLEARLY IN THE
LATERAL SENSORY CLUSTERS.
-!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
-!- SIMILARITY: Contains 1 Sema domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.sib-sib.ch/announce/
or send an email to license@sib-sib.ch).
EMBL; L26082; AAA88789.1;

```

```
EMBL: AE003621; AAF52695.1; ..
FlyBase: FBgn0011259; Sema-1a.
InterPro: IPR003659; Plexin-like.
InterPro: IPR002165; Plexin_repeat.
Pfam: PF01437; PSI; 1.
Pfam: PF01403; Sema; 1.
SMART: SM00423; PSI; 1.
SMART: SM00423; Sema; 1.
Signal: Developmental protein; Transmembrane; Glycoprotein;
Neurogenesis.
SIGNAL 1 19 POTENTIAL
CHAIN 20 771 SEMAPHORIN 1A.
DOMAIN 21 606 EXTRACELLULAR (POTENTIAL).
TRANSMEM 609 629 POTENTIAL.
DOMAIN 630 771 CYTOPLASMIC (POTENTIAL).
DOMAIN 630 771 SEMA.
DOMAIN 689 694 POLY-PRO.
CARBOHYD 41 41 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).
SEQUENCE 771 AA; 86935 MW; 752E7D0516F33D45 CRC64;

Query Match 17.1%; Score 932.5; DB 1; Length 772;
Best Local Similarity 35.0%; Pred. No. 2.1e-52;
Matches 219; Conservative 156; Mismatches 233; Indels 87; Gaps 20;

67 TLVIAARGHVYVDITSTHTSEIVSGKLTWKSRAVDTCRNGKNGKDECHNFIKVLK 126
68 .....HCLVQQRLVWTSPEDDTKACLVKNGKSEAQNYIRVWV 111
69 SLLIGARNTVFHLS-----HCLVQQRLVWTSPEDDTKACLVKNGKSEAQNYIRVWV 111
127 KNDALFVCGTNAFNSCRNYKHC-----TLSPFGDPSPGSMARCPYDAKHANVLPADSKL 182
128 .....KNDALFVCGTNAFNSCRNYKHC-----TLSPFGDPSPGSMARCPYDAKHANVLPADSKL 182
112 PPGRLFGVGTNSFRPMCHTYIISDSNVTLA-----TKGQAVCPYDPHNSVSLADNEL 168
163 YSATVTDFALDAVYRSGESPTLRTVNDKSKLKERFYQAVDYGDIYVPPFRIEAVE 242
164 .....YSATVTDFALDAVYRSGESPTLRTVNDKSKLKERFYQAVDYGDIYVPPFRIEAVE 242
169 YSGTVADFGSDPFIYRE-----FQTEQYDLSLNATNFVSSTFGDPVVFPPFETAVE 223
243 YNTMGAVFPRVAQCKNKGSCQVLEKCVTSFLKARLNGVSPGDSHFYFNLCAVTCV 302
244 FVNSKAVSRVARCKMKGGEFR-FRNRWTSFLKSRNLGSDYFDFYFNEISASNL 282
303 IR-----IMGRDYVATFTPYNSIPGSAVCAYDNALIASVFTGFEKQKSTSTWTFYED 358
283 VEGQYGSMSKLIYGVNTPNSIPGSAVCAYDNALIASVFTGFEKQKSTSTWTFYED 358
359 ERYPRPCCGAGSSS:RYATSNRPPDTLNTKTHFLMRDAVSSINRPFVFTMPRY 418
360 .....ERYPRPCCGAGSSS:RYATSNRPPDTLNTKTHFLMRDAVSSINRPFVFTMPRY 418
363 AKVPDPFGSC-----HNFSFALPDPFLNFKTHSLNENVFAFSPQFLARTSTY 394
419 RTTKIAVD---TAAGFYGNHTVAFILGSEKGIILAFI-ARIGNSGFINDSLFLBMSVYN 473
395 RTQIAVDQIKTPG--KTYDVIVGTDHGLIKIYNNRESADSKVTSVVISIDVLT 452
474 SKKCYGVGVEDKRIIMGMQLDRASSLSY-----VAFSTCVIKVPLGRG--ERHSGKQKT 524
453 KS-----EPIRNLIVATMQVDQPDSSYDQGLIIVTDSQVVAIQHRCNDKITSQSR 507
525 CIASRDFYGVWKEGAC-SHLSFN---SLTTEQDIERGNTDGLGCHNSFVALNGHSS 561
508 CVALQDFYCAWQKIAKCRSHGAPRLEENFYQNVATCG-----HAAC 551
582 LIPSTTSSTAQGVESGGLWKKHLQSDSTDPGLGAVSHSHQDKGVIRSEYILKG 641
552 PGKINSKANKAGEQKGFNDM-----CILCS-----RRQSKQDQEIIONDKN 594
642 HQQLVPVTLIAIVLAFVNGAVFS 666
595 FDIINAQYVTEILWAVLAGSIFS 613
```

```
RESULT 11
SN3A_RAT ID SN3A_RAT STANDARD: PRT: 772 AA.
AC Q63548;
DT 30-MAY-2000 (Rel. 39, Created);
DT 30-MAY-2000 (Rel. 39, Last sequence update);
DT 15-SEP-2003 (Rel. 42, Last annotation update);
DE Semaphorin 3A precursor (semaphorin III) (Sema III).
GN SEM3A.
OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC NCBI_TaxID=10116;
RN 1.
PP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RA MEDLINE=9707089; PubMed=8915827;
RX Giger R.J., Wolfer D.P., De Wit G.M.J., Verhaeger J.:
RT "Anatomy of rat semaphorin III/collapsin-1 mRNA expression and
PT relationship to developing nerve tracts during neuroembryogenesis.";
RL J. Comp. Neurol. 375:378-392(1996).
CC 1- FUNCTION: MAY BE INVOLVED IN GUIDING GROWING AXONS TOWARDS THEIR
CC TARGETS BY FORMING A MOLECULAR BOUNDARY THAT INSTRUCTS AXONS TO
CC ENGAGE IN THE FORMATION OF SPECIFIC NERVE TRACTS. BINDS TO
CC NEUROFILIN.
CC 1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC 1- DEVELOPMENTAL STAGE: AT E-11, EXPRESSION WAS RESTRICTED TO THE
CC OLFACTORY PIT, THE BASAL AND ROSTRAL SURFACE OF THE TELECEPHALIC
CC VESICLE, THE EYE ANLAGE, THE EPITHELIUM OF RATHKE'S POUCH, AND
CC SOVITES. AT LATER DEVELOPMENTAL STAGES, IT WAS WIDELY DISTRIBUTED
CC IN NEURONAL AS WELL AS IN MESENCHYMAL AND EPITHELIAL STRUCTURES
CC OUTSIDE THE NERVOUS SYSTEM. AFTER BIRTH, MESENCHYMAL LEVELS
CC DECREASED RAPIDLY AND EXPRESSION BECAME RESTRICTED TO SPECIFIC
CC MITAL CELLS, NEURONS OF THE ACCESSORY BULB AND CEREBRAL CORTEX,
CC CEREBELLAR PURKINJE CELLS, AS WELL AS A SUBSET OF CRANIAL AND
CC SIGNAL MOTONEURONS.
CC 1- DOMAIN: STRONG BINDING TO NEUROFILIN IS MEDIATED BY THE CARBOXY
CC THIRD OF THE PROTEIN.
CC 1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC 1- SIMILARITY: Contains 1 Sema domain.
CC 1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL databank -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. See http://www.isb-sib.ch/annouce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X95286; CAA64607.1; ..
CC InterPro: IPR007110; Ig-like.
CC InterPro: IPR003599; Ig.
CC InterPro: IPR003026; Ig_XHC.
CC InterPro: IPR003659; Flexin-like.
CC InterPro: IPR001627; Sema.
CC Pfam: PF00047; Ig; 1.
CC Pfam: PF01403; Sema; 1.
CC SMART: SM00429; IG; 1.
CC SMART: SM00423; PSI; 1.
CC SMART: SM00423; Sema; 1.
CC SMART: SM00630; Sema; 1.
CC PROSITE: PS00835; IG-LIKE; 1.
CC Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
CC Developmental protein; Glycoprotein.
CC SIGNAL 1 220 POTENTIAL.
CC CHAIN 21 772 SEMAPHORIN 3A.
CC DOMAIN 240 538 SEMA.
CC DOMAIN 577 665 IG-LIKE C2-TYPE.
CC DOMAIN 728 770 ARG/LYS-RICH (BASIC).
CC DISULFID 650 723 BY SIMILARITY.
```

```

T CARBOHYD 53 53 N-LINKED (GLCNAC... (POTENTIAL)
T CARBOHYD 125 125 N-LINKED (GLCNAC... (POTENTIAL)
T CARBOHYD 591 591 N-LINKED (GLCNAC... (POTENTIAL)
Q SEQUENCE 772 AA: 88808 MW; 2409076127F9F2D2 CRC64;

Query Match 15.78; Score 856; DB 1; Length 772;
Best Local similarity 32.38; Pred. NC 1.3e-47;
Matches 209; Conservative 104; Mismatches 25; Indels 54; Gaps 20;

47 GRNTQRRHLDICVIMINNGT-----LVIAARDHVTVEIDT 83
25 GKNNVPLKLSYKEMLSNNVITNGANSSVHTFLDEERSPLVGCARDHFSNLS-- 82
84 SHREELVCSKRLTKWKSQADVCTCRNGKH-KHECHNFIVLLKLNKDDALFVCGTNAEXP 142
83 ---VNKDFGKIVFVSVYTRDECKWAGKILKFCACNTIKVLKAYHOTHLYACGTGAERP 139
143 SCR-----NVKVCOTLEPFGEDEF--SGWARGPYAKAKHVALFADGKLYSATVDFLADA 195
140 ICYVIEVCHREERNIFLQSHSHENGKSGPYFKJLTASLLLDGSHVSTADFPKQDP 195
196 VVRSLOESPLTVKHSKWLKEPYVQA-----VQYGDYIVFFREIAVEVNTMGK 248
200 AIPETLGHRIPIREFQDSKMLNDPRISAHILIPESNDPEDEKVVYFFRANADGHSOK 253
249 VVPRVACVKNMGSGGQRYLEKWTSLFARLNCVFG--DEHYFNILQAVTDYIN 304
260 ATVARIGCIKNDFFGG-HRSVNRKWTFLKARIICSVFGRNGLDTHF-----DELDQVFL 311
305 INGRD---VYLAFTSTPVSIRHSACVAYEMLIASVFTGREKCKSPDSTVTPQSR 360
314 MNSKQKPPVYGVFTSSNIFASVCMYKSDRVFLGPIAHREGNYQWVF-QSR 372
361 VPKRPCCAGSSGLRYASNEFFDTLFIKTHRMDSAVSIFNRWFLFIMVYR 420
373 VVYPRPQCTP-SKTEGDFDSTKQLEDDVITFANSHFPMVPEPINKRELYKNTDNYQF 431
421 TKIADPTAGPYONHTVYVVGSEGIILKFAILKSGFINDSLFLENSVNSKCSYD 430
432 TQIVDVRVDAEDQGVYVFGIDVGTVLKVVSPKRWEDLEVLVLEEMVYR----- 434
491 GVDEKRIYNGQLDRASSLSLVAFSTCVIKYPLGRCERHGRCKKXTCIASRDEYCGWKEGG 540
465 --EPTTISAMELSTKQQLYIGTAGVAQPLHRCQYGRACRECLARDPYCAW--DSS 540
541 ACHLSLSPNS-RLTFEODERGNTGGLCC-----HNSFVALNGHS--SSLLPSTTTSDST 592
541 SCRYFPPTAKRTFRRCQIRNG--DPLTHCSDLOHICNH---HGHSLERKIIYGVENSSTF 595
593 AQEGYSGRGYLDWKH-LDSPDSFDPLGAVSSHHCQKQKGVIESVLK 640
596 LECSPKQALVYQVQFRRNEDRKEEI-RVGDHRTTEQGLLRSIQCK 642

```

## RESULT 12

M3A MOUSE

```

C SW3A MOUSE STANDARD; PRT; 772 AA.
C Q08685; Q62180; Q62215;
T 30-MAY-2003 (Rel. 39, Created)
T 30-MAY-2003 (Rel. 39, Last sequence update)
T 15-SEP-2003 (Rel. 42, Last annotation update)
E Semaphorin 3A precursor (Semaphorin III; (Sema III; (Semaphorin D'
E (Sema D)
N SEMA3A OR SEMAD OR SEMD.
S Mus musculus (Mouse).
S Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
X Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Mus.
X NCBI_TaxID=10090;
X [1]
N SEQUENCE FROM N.A.
P S-RAIN=NMRI; TISSUE=Embryo;
X MEDLINE=95267431; PubMed=7748561;
X Pleschel A.W., Adams R.H., Betz H.;

```

```

RT "Murine semaphorin D/collapsin is a member of a diverse gene family
RT and creates domains inhibitory for axonal extension.";
RL Neuron 14:941-948(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97473885; PubMed=9331345;
RA Taniguchi M., Yuasa S., Fujisawa H., Naruse I., Saga S., Mishina M.,
RA Yagi T.;
RT "Disruption of semaphorin III/D gene causes severe abnormality in
RT peripheral nerve projection.";
RL Neuron 19:519-530(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX Kimura T., Fishman M.C.;
RT "cDNA sequence of mouse collapsin/semaphorin III.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 107-772 FROM N.A.
RX "SSEA fetal brain;
RX MEDLINE=95267432; PubMed=7749562;
RA Vessier-Smith E.K., Leonardo E.D., Shatz C.J., Tessier-Lavigne M.,
RA Goodman C.S., Kolodkin A.L.;
RT "Semaphorin III can function as a selective chemorepellent to pattern
RT sensory projections in the spinal cord.";
RL Neuron 14:549-558(1995).
CC -!- FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE. MAY FUNCTION TO
CC PATTERN SENSORY PROJECTIONS BY SELECTIVELY REPELLING AXONS THAT
CC NORMALLY TERMINATE DORSALLY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN EMBRYONIC DEVELOPMENT
CC -!- IN DISTINCT REGIONS OF THE NEUROECTODERM AND MESODERM.
CC EXPRESSION BECAME MORE EXTENSIVE AT LATER STAGES.
CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
CC THIRD OF THE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X85993; CAA59385.1;
CC EMBL: D85028; BAA19793.1;
CC EMBL: A41541; AA177612.1;
CC EMBL: J40484; AAA73934.1;
CC PIR: I48747; I48747.
CC PIR: I58163; I58169.
CC MGD: MGI:137558; Sema3a.
CC InterPro: IPR007110; IG-like.
CC InterPro: IPR003599; IG.
CC InterPro: IPR03006; IG_MHC.
CC InterPro: IPR03659; Plexin-like.
CC Pfam: PF01403; Sema; 1.
CC Pfam: PF00047; IG; 1.
CC SMART: SM0409; IG; 1.
CC SMART: SM0423; PSI; 1.
CC SMART: SM00630; Sema; 1.
CC PROSITE: PS00835; IG LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental Protein; Glycoprotein.
FT SIGNAL 1..20 POTENTIAL.
FT CHAIN 21..772 SEMAPHORIN 3A.
FT DOMAIN 240..539 SEMA.
FT DOMAIN 579..665 IG-LIKE C2-TYPE.
FT DOMAIN 728..773 ARG/LYS-RICH (BASIC).
FT DISCID 650 723 BY SIMILARITY.
FT CARSDYD 53 53 N-LINKED (GLCNAC... (POTENTIAL).

```

```

T CARBOHYD 123 125 N-LINKED (GLCNAC... (POTENTIAL).
T CARBOHYD 591 593 N-LINKED (GLCNAC... (POTENTIAL).
T CONFLICT 193 193 D -> N (IN REF. 4).
T CONFLICT 207 207 H -> D (IN REF. 1).
T CONFLICT 253 253 D -> G (IN REF. 1).
T CONFLICT 352 352 F -> G (IN REF. 4).
T CONFLICT 403 403 A -> G (IN REF. 1).
T CONFLICT 571 572 OH -> ED (IN REF. 1).
T CONFLICT 616 620 EDRKE -> RRKR (IN REF. 1).
T CONFLICT 623 623 P -> K (IN REF. 4).
T SEQUENCE 772 AA; 98799 MW; E69A0852B5-0A8C3 CRC64;

Query Match 15.7%; Score 253; DB 1; Length 772;
Best Local Similarity 32.1%; Pred. No. 2,7e+47;
Matches 208; Conservative 102; Mismatches 257; Indels 78; Gaps 18;

Y 47 GRNTTQRRLDQIMINMGT-----LYIAARDHIYTVDT 93
| | | | | | | | | | | | | | | | | | | | | |
D 25 GKQVRLKLSYKEMLESKNVITFNGLANSSVHTLLDERSRLVVGAKHIFSFNL-- 92
| | | | | | | | | | | | | | | | | | | | | |
Y 84 SHTEEYCSKKLTKNSROADYDTCRKGKH-KDECHNFIKVLKNDLAFVCGTNAFNP 142
| | | | | | | | | | | | | | | | | | | | | |
b 83 ---VNIKDFQKIVWPVSYYTRRDECKWAGSDILKECANFIKVLKVNQTHLYACGTGAPRP 139
| | | | | | | | | | | | | | | | | | | | | |
Y 143 SCR-----NYKXDTLEPPGDEF--SQMASCPYDAKHANVALFADGKLYSATVTDPLAIDA 195
| | | | | | | | | | | | | | | | | | | | | |
D 140 ICTYIVGHHPEDNIFKLGDSHFNGRSGSPDPKLLTASLLIDELYSGTANDWGRDP 199
| | | | | | | | | | | | | | | | | | | | | |
Y 196 VYRSLGESPTLRTVXHSKWLKEPYEVA-----VQYGYIYFFRERAEVNTMOK 249
| | | | | | | | | | | | | | | | | | | | | |
D 200 AIFRTLGHHPITEQHSKWLNDPFIHAHLIPEDCNEDPKVWFFPENAIDGHSKK 259
| | | | | | | | | | | | | | | | | | | | | |
Y 243 VVFPVAVQCKMDGMSORVLEKQWTSFKARLNCVFG-----DSHFYENILCAVTDYR 304
| | | | | | | | | | | | | | | | | | | | | |
D 260 ACHARIGQCKNDPGG-HRSIYKWTFTFKARLICSVPGRSGDTHF-----DEGCVEL 313
| | | | | | | | | | | | | | | | | | | | | |
Y 305 INGRQ-----VVIATFSTPNISIPGSAVCAYCLDASVFTGRFKECKSPDSTWTPVPER 360
| | | | | | | | | | | | | | | | | | | | | |
b 314 MNSKDPKNIVYGVFTSSNIFKSAVMSYKSDVRVELGFAHEDGNYQWVPY-QGR 372
| | | | | | | | | | | | | | | | | | | | | |
Y 361 VPXPRGCGAGSSSLERYATSNFPDCTLNFITKTHPLMDAEVPSFNRWFRTWYVRL 420
| | | | | | | | | | | | | | | | | | | | | |
b 373 VPYPREGTGP-SKTFGFGDSTKLPDEVITEARSHAPXANVPVFNRRIM-KTDWYGF 431
| | | | | | | | | | | | | | | | | | | | | |
Y 421 TXAVDTAGPYGNHVTPLGSKGILLKPLRIGNSGFLNLSPLSESVYNSEKCSVD 480
| | | | | | | | | | | | | | | | | | | | | |
D 432 TQVWDVDAEDCCYDMFTGTDVGVLKVVSVFKETWHCLSEVILEXTVPR----- 484
| | | | | | | | | | | | | | | | | | | | | |
Y 481 GVECKRMGMQLDRASSSYVAFTSTVYKVPACRERHKKKKTCTIASRDPYCGNKEGG 540
| | | | | | | | | | | | | | | | | | | | | |
D 485 --EPTTISAMELSTRCCCLYIGSTAGVACILPHRCDIYKKAACGLARCPYCAN--DGS 540
| | | | | | | | | | | | | | | | | | | | | |
Y 541 ACSHUSPNS-PUTTEGDIERGNTGGLGDSHNEFVALNGHSSS-----JSTTTTSSTAGE 595
| | | | | | | | | | | | | | | | | | | | | |
D 541 SCRYFPPTAKRRTRRCIRNG--DPLTHCSOLQHDHWHGSPSEERIIYGVENSSTFIFC 598
| | | | | | | | | | | | | | | | | | | | | |
Y 596 GYSGRGMLDKWILLDSPSTDFLGAVSSNHQKGVTPRESIK 640
| | | | | | | | | | | | | | | | | | | | | |
b 599 SPKSQALVYVQFQRNEDRKEE--RYGDHITREGGELINSQK 642
| | | | | | | | | | | | | | | | | | | | | |

RESULT 13
M3A_HUMAN
D_SNA3A_HUMAN STANDARD; PRT; 771 AA.
C Q14563;
T T 30-MAY-2000 (Rel. 39, Created:
T T 30-MAY-2000 (Rel. 39, Last sequence update:
T T 15-SEP-2003 (Rel. 42, Last annotation update:
E Semaphorin 3A precursor (Semaphorin I.1) (Sema I.1).
S SNA3A.
S Homo sapiens (Human).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Query Match 15.6%; Score 852.5; DB 1; Length 771;
Best Local Similarity 32.1%; Pred. No. 2,9e+47;
Matches 208; Conservative 106; Mismatches 259; Indels 83; Gaps 20;

QY 44 HKPGRNTQRRLDQIMINMGT-----LYIAARDHIYTVDT 80
| | | | | | | | | | | | | | | | | | | | | |
```

```

b 22 YQCKXNVPELKLKSYKEMLENNVITFNGLANSSVATFLDZERSELYVGAKDHIESD 81
b 81 IDTSHTTEIYCSKLLTKMSQADVCTORMKGR-KOCHNFIKVLKQDDA-FVCGTWA 129
b 82 L-----VNIKDPQKIVMPVSVTRDECKWAGKDLKSCANFIKVLKAYNCTHLYACGTGA 136
b 140 ENRSCR-----NKKMUTLEFGDSF--SGMARCEYAKAHANVALFADGKLYSATVTDFLA 192
b 137 FPHICTYRIGHPEPEDIKLESHFENGKSGSPYDKLTAAGLLDGLGYSTAADEMG 196
b 193 IDAVIRSLGESPLTRVKKDSKALKEPVPQCA-----VOYGVYVFFFRGIAEYNT 245
b 197 RDAIFRPLGHHHPIRTECHDSRLNDPKLSAHLSESSENPEDUKVYFFFRNAIDGHE 256
b 246 MGKVVPFPAQVCKNNDGSGQVVEKQWTFELKARLNCVFG-----DSHFYFNLQAVTD 301
b 257 SGKATHARIQICKNDFGG-HRSJUNKWTFELKARLICSVEGPNKGDTHF-----DELQD 310
b 302 VIR-NGRD-----VVIATFSPYKSIPOGSAVCADYLDIASVFTGRFKEQKSPDSTNPF 357
b 311 VFLKNFQPNFVYGVFTSSNIFKGSVCYKXNSDYRVRVFLQPAHRDGPYQWVPY 369
b 359 DERVPKPRQCCAGSSSERYATSNPEPDTIPIKTHPYMDRAVESIFKPFMFRTMR 417
b 370 QGRVYFEPGTCF-SKTFGGFDSKLPDQVITFAPSHPAVYVPPFPMNRPVINTDN 426
b 418 YRLTKIANDTAAQPYQNTVYVLSSEKGIILKFLARIGNSGFLNDSLFLEKMYNYSKC 477
b 429 YQFTQVYVDAEDGGYDYNFISTDGTGTVKVSIFKRTYGLRVLSEMYVFP-----484
b 478 SYDGVEDKRMGMCLDRASSLYVAFSTCVIKVPLGRCEHKGKCKTCTIASRDPYGNIK 537
b 485 -----EFTASAMELSTKQQLYIGSTAGVAQPLHRCDIYKACAECCCLAPPYCAM-- 537
b 538 EGGACSHLSPNS-ELTFEQDIERANIDELQD-----HNSFWALNHS--SALLPSTTSD 592
b 538 DGSACSRYPFTAKRTRRQDINQ--DELTHCSLHEDN--HGRSFEBEIIYGVENSS 592
b 591 STAGEYESGGMLDAKHLDSPDSTUPLGAVSHNHQKKVIRE 637
b 593 TFLCSPKSGACVYQWQPRNEERKEEIVRVDHIIRTDGGLLRS 638

```

## RESULT 14

```

SMIA_TRIOP
c SMIA_TRIOP STANDARD; PRT; 712 AA.
c Q26972;
c 30-MAY-2000 (Rel. 39, Created);
c 30-MAY-2000 (Rel. 39, Last sequence update);
c 16-OCT-2000 (Rel. 40, Last annotation update)
c Semaphorin-1A precursor (Semaphorin-1).
c SEMA-1A OR SEMA-1.
c Tribolium confusum (Confused flour beetle).
c Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
c Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
c Tenebrionidae; Tribolium.
c NCBI_TaxID=7071;
c (1).
c SEQUENCE FROM N.A.
c TISSUE=Embryo;
c MEDLINE=94094332; PubMed=8269517;
c Kelodkin A.L., Matthes C.J., Goodran C.S.;
c "The semaphorin genes encode a family of transmembrane and secreted
c growth cone guidance molecules.";
c Cell 75:1389-1399(1993).
c -- FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE.
c -- SUBCELLULAR LOCATION: Type I membrane protein.
c -- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
c -- SIMILARITY: Contains 1 sema domain.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC -----
CC EMBL: L26980; AAA16609.1;
CC InterPro: IPR003659; Plexin-like.
CC InterPro: IPR002165; Plexin_repeat.
CC InterPro: IPR001627; Sema.
CC Pfam: PF01437; PSI; 1.
CC Pfam: PF01403; Sema; 1.
CC SMART: SMC423; PSI; 1.
CC SMART: SMC630; Sema; 1.
KW Signal; Developmental protein; Transmembrane; Glycoprotein;
KW Neurogenesis.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 712 SEMAPHORIN 1A.
FT DOMAIN 21 631 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 602 622 POTENTIAL.
FT DOMAIN 623 712 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 213 536 SEMA.
FT DOMAIN 473 473 POLY-VAL.
FT DOMAIN 612 616 POLY-VAL.
FT CARBHYD 42 42 N-LINKED (GLCNAC) (POTENTIAL).
FT CARBHYD 69 69 N-LINKED (GLCNAC) (POTENTIAL).
FT CARBHYD 161 161 N-LINKED (GLCNAC) (POTENTIAL).
FT CARBHYD 265 265 N-LINKED (GLCNAC) (POTENTIAL).
SQ SEQUENCE 712 AA; 79751 MW; C5734FESB9C09FE3 CRC64;
Query Watch 15.5%; Score 845; DB 1; Length 712;
Best Local Similarity 28.9%; Pred. No. 40-47;
Matches 230; Conservative 130; Mismatches 270; Indels 166; Gaps 30;

```

```

QY 12 LHPAGAGFPEDSEF-----SIHGNTYKQYVFGVGHKPGQNTTORHLDIQVIMNK- 65
DB 12 LIALCHAMPDSSSKLINFKPSVSKSPT-----GNATFPCH-----FIVLQK 54
QY 66 --GTLYIARCHIYVDI-DTSHTEIYCSKLLTKMSQADVCTORMKGRKDECHNEIK 122
DA 55 DETSLVGGRKRVKNSIFTSERK-----GGRIVFSSDANGCLCAGKTDCCQNTIR 110
QY 123 VLLKXNDALFVCGTNAPNFCENYKMTLEPFD-BFSGVACFPYDAKHANVALFADGK 181
DB 111 LLYSSEPKLVICGNSYKPLCRTVAFKEGKYLVEKEVEGIGLCFPYNEHNSSTSVYNGQ 170
QY 182 LYSATVTDPLADAVYRSLGESPLTRVKKDSKALKEPVPQAVCYCYIYFFPREAY 241
DB 171 LFSATVADFGSDPLIYRE-----PQRTSLDLKQLNAPNFVNSVAYCYIFFFYRETA 225
QY 242 EYNTMGKVVPFPAQVCKNDMGSGSRVLEKQWTFELKARLNCVPGDSHFYFNILQAVTD 301
DB 226 EYNGCGVIYRVARVCKDCKGGPHQSDR-WTSLKARLNCVPGDYFFYFDEIQS-SD 294
QY 302 VI--RINGRD---VVLATFSTPYNSIPGSAVCADYLDIASVFTGRFKEQKSPDSTWTPV 356
DB 285 IVEGRVNSDDSKLIYG-LITPVNNAIGSA-CAYQADILAVFEGSPKQETINSNKLVP 344
QY 357 PDERVPKPRQCCAGSSSERYATSNPEPDTIPIKTHPLMDRAVPSIFRPFWRTRYV 416
DB 345 PQNLVPEPRPGQCVRSR-----LPDKNVNFIKTHSLMED-VPAJFGKPLVRLVSL 395
QY 417 RYRLTKIAVCTAAGPYQNH--TVVFLGSEKGIILKFLARIGNSGFLNDSLFLEMSVNS 474
DB 396 QRFRTAITVDPQVKTINQYLVLVIGTDDGVK----- 430
QY 475 EKCSYDGVEDKRMGMCLDRASSLYVAFSTCV-----IKVP----- 511
DB 431 -----AVNIPKRAKALLRYKRTSVHPHGPVQLKIAPGYGVVVVVGKDEI 478
QY 512 ----LGRCEHKGKCKTCTIASRDPYCGWIKEGGASHLSPNSRQTF-EQIERGN----- 561
DB 479 RLANNHCASKTRC-KDCVCELODPCAMDANQLCVSIDTIVTSYRFLICDVVVRGDCNCK 537

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licenses@isb-sib.ch).

562 ---TDLGSECHNEFVALNGSSLLPSTTSSTAGCTGTSAGVLDKWLHLLSSDSTP 618  
538 SGTQKXVTKK-----PSEVENET-----NSIDKEL--DSSTP 572  
619 LGAVSHHQKKGVIRESYKQ---HDCVPPV---LALAVLAPWGA---VSGITV 670  
573 LKLTGLDSDSDCPV-SENSIGCAVQQLVIVTAGTLHVVVAVSIVGLFWYSGLSV 631  
671 YCV--CCHARKOVAVVCRKE--KELTHSRGSM-SVTK:SGLFGTQKQKPPAILTP 725  
632 FAXHSDSCYPEAFTEQNHLEJLSANQCYLTTPRANRAVHLVWVSSSTPRPKXNLD 691  
726 LMHNGKLCATPGNTAKY 74:  
692 VSKDNIASDGTQK: 707

RESULT 15  
ZIB BRARE  
D S21B BRARE STANDARD; PRC: 778 AA.  
C QM686;  
T 30-XAY-2000 (Rel. 39, Created:  
T 30-XAY-2000 (Rel. 39, Last sequence update)  
T 15-SEP-2003 (Rel. 42, Last annotation update)  
E Semaphorin ZIB precursor (Semaphorin 1B) (Sema-21B).  
N SEMA21B OR SEMA3A5  
S Brachydanio rerio (Zebrafish) (Danio rerio).  
C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
C Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
C Cyprinidae; Danio.  
N NCBI\_TaxID=7955;  
X [1]  
P SEQUENCE FROM N.A.  
X MEDLINE:99425174; PubMed:10495275;  
X Roos M., Schachner M., Bernhardt R.R.,  
T "Zebrafish semaphorin zib inhibits growing motor axons in vivo.",  
X Mech. Dev. 87:103-117(1999).  
C -!- FUNCTION: NIGHT NORMALLY INFLUENCE THE MIDSEGMENTAL PATHWAY CHOICE  
C OF THE VENTRALLY EXTENDING MOTOR AXONS BY CONTRIBUTING TO A  
C REPUSIVE DOMAIN IN THE POSTERIOR SOMITE.  
C -!- SUBCELLULAR LOCATION: Secreted (by similarity).  
C -!- TISSUE SPECIFICITY: EXPRESSED IN RHOMBOMERES THREE AND FIVE, AND  
C IN THE POSTERIOR HALF OF NEWLY FORMED SOMITES WHICH IS AVOIDED BY  
C VENTRALLY EXTENDING MOTOR AXONS.  
C -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.  
C -!- SIMILARITY: Contains 1 sema domain.  
C -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
C This SWISS-PROT entry is copyright. It is produced through a collaboration  
C between the Swiss Institute of Bioinformatics and the EMBL outstation  
C the European Bioinformatics Institute. There are no restrictions on its  
C use by non-profit institutions as long as its content is in no way  
C modified and this statement is not removed. Usage by and for commercial  
C entities requires a license agreement (See <http://www.isb.ch/announce/>  
C or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
C  
C EXBL: AF063382; AA028101.1;  
C ZFIN: ZDB-GENE-991209-6; semaph.  
C InterPro: IPR007110; Ig-like.  
C InterPro: IPR003593; Ig.  
C InterPro: IPR033006; Ig\_YHC.  
C InterPro: IPR003659; plexin-like.  
C InterPro: IPR002165; plexin\_repeat.  
C InterPro: IPR001627; Sema.  
C Pfam: PF00047; Ig.  
C Pfam: PF01437; PSI.  
C Pfam: PF01403; Sema.  
C SMART: SM00409; IG.  
C SMART: SM00423; PSI.  
C SMART: SMC0630; Sema.  
C PROSITE: PSS0835; IG\_LIKE.  
C Signal: Immunoglobulin domain; Multigene family; Neurogenesis;  
C Developmental protein; Glycoprotein.

FT SIGNAL 1 17  
FT CHAIN 18 778  
FT DOMAIN 24: 539  
FT DOMAIN 579 668  
FT DOMAIN 721 776  
FT DISULFID 652 716  
FT CARBOHYD 54 54  
FT CARBOHYD 127 127  
FT CARBOHYD 593 593  
SC SEQUENCE 778 AA; 59304 MW; 4D36F432AE21995 CRC64;  
POTENTIAL.  
SEMAPHORIN ZIB.  
IG-LIKE C2-TYPE.  
ARG/LYS-RICH (BASIC).  
BY SIMILARITY.  
N-LINKED (GLCNAC). (POTENTIAL).  
N-LINKED (GLCNAC). (POTENTIAL).  
N-LINKED (GLCNAC). (POTENTIAL).  
Query Match: 15.5%; Score 845; DE 1; Length 778;  
Best Local Similarity 34.3%; Pred. No. 9e-47;  
Matches 194; Conservative 94; Mismatches 204; Indels 74; Gaps 17;  
QY 66 GTLYIARQHIYTVD-IDTSHTBIYCSKLLTWKSRADYDTCKMKKH-KDCENIKV 123  
DB 68 GRLEVGAKDHVLSFNLDVINDMQQL-----SWPSPSPRRDECKKAGKGVQKECANIKV 122  
QY 124 LKKNDDCALFVCGTNAENFSCNYKM-----DTLEPEGDEF-SGVARCPYDAKHANVALF 177  
DB 123 LQPNQTHLVACGTGAFHPVCAHVEVSKRSEDNTFRUGSSFENGGRGKSPYDPKQOTASML 182  
QY 178 ADGLYSATVTDFLAIDAVIYRSIGESPT-RTVXHSKWLKPEYFV-----CAVDY 228  
DB 183 IDGELYAGTSADFMGRDPFAIFPTLGKXHPITRTEQHSRWLNDPRFVSVHL-PESQNAED- 241  
QY 229 GDYIYFPREIAVYNNYKGVFFRVAVQVCKNDMGSGQORVLEKQNTSFLKARLNCSPG- 287  
DB 242 -DKYLFPRFENAI DGEQISKATHARIQGLCKNDFFGG-HRSLVNKWT-FLKARVCSVPGL 299  
QY 288 ---DSHEVENIQAQVTDVIRINGD---VVLATFSTPYNSIPGSAVCAVCMLEIASVET 340  
DB 320 NGIDTHP-----DELQDVFLLXSSKDPKNPIIYAVFTTSSK-FKGSAVCYSMADIRRVFL 354  
QY 341 GRPKEQKSPDSTWTPVPRDERVFKPRPGCCAGSSSLERYATSNPEPDDTLNFKTHPLMDE 400  
DB 355 GPVAHROGPNVQWYFPLK-RVPIYPRPGTGP-SKFDGFESTKDFPDVITFARSHFAMYN 412  
QY 421 AVPSIPRPPWELRTVRYRLTKIADVTAAGYQNHVVVLOSEKGIILKFLARIGNSGFL 460  
DB 413 PVFFPNNHPIIKTCDVDYQFTQIVVDVREADGGVYDMFICTDMGTQVKKVSIPEGTHD 472  
QY 481 NDSJFLEEMSYVNSEKSYDGVDEKRMGMOLDRASSSLVVAFTCVIKVPLGRCERHGK 520  
DB 473 LEEVLEEMTVFR-----EPTATTCAMEJSTKQQQJYLGSAIGVSQKPLHRCDDVYVK 523  
QY 521 CKKTICIASRDPYCGWIKEGGACSHLSPNS-RLTPEQDIERTGNTDGLGDCNKSFYALNGHS 579  
DB 574 ACABECLAKDQYCAW--DGSQCSRYFP-ATKBTBTRQDIANGD----- 563  
QY 580 SLLPSTTTSSTAGQGVESRGGMLD 605  
DB 564 ---PLTQCSDLQHEDEADGEAGLLD 585

Search completed: October 23, 2003, 17:09:37  
Job time : 22 secs



```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

protein - protein search, using sw model

on on:      October 23, 2003, 17:09:37 , Search time 37 Seconds
           (without alignment)
           4653.056 Million Cells Updates/sec

file:       US-09-856-681-2
rfc score:  5450
sequence:   1 KRSEALLSYFLHIFAGAF.....PKPFAFLSTSHKPNDACT 1033

oring table: BLOSUM62
            Gapop 10.0 , Gapext 0.5

arched:     83525 segs, 253252604 residues

total number of hits satisfying chosen parameters: 83525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database:     SPTREMBL_231.*

```

**tabase :**

1: sp\_arhea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebra.\*  
6: sp\_mammal.\*  
7: sp\_mbc.\*  
8: sp\_opanarello.\*  
9: sp\_plant.\*  
10: sp\_plant.\*  
11: sp\_plant.\*  
12: sp\_plant.\*  
13: sp\_plant.\*  
14: sp\_plant.\*  
15: sp\_plant.\*  
16: sp\_plant.\*  
17: sp\_plant.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SL-2003-ES

SUIT No.	Query Match	Score	Length	DB	ID	Description
1	100.0	5450	1030	4	Q9H2F6	Q9H2E6 homo sapien
2	99.7	5431.5	1049	4	Q9P2H9	Q9P2H9 homo sapien
3	92.2	5024.5	1005	11	Q9EQ71	Q9EQ71 mus musculus
4	67.5	3677	693	4	Q9ESW4	Q9ESW4 homo sapien
5	54.7	2980.5	574	4	Q9ESN8	Q9ESN8 homo sapien
6	54.4	2963	587	11	Q8BQ10	Q8BQ10 mus musculus
7	54.3	2961	605	11	Q8BXZ7	Q8BXZ7 mus musculus
8	54.2	2952	562	4	Q9G5V4	Q9G5V4 homo sapien
9	29.7	2947	562	4	Q8NC49	Q8NC49 homo sapien
10	26.2	2621.5	507	4	Q8ETC4	Q8ETC4 homo sapien
11	33.68	48.1	1073	4	Q8NFY4	Q8NFY4 homo sapien
12	33.49	43.2	1327	4	Q8NRP5	Q8NRP5 homo sapien
13	23.15	42.5	998	4	Q8NRP6	Q8NRP6 homo sapien
14	23.14	42.5	1011	4	Q8NRP3	Q8NRP3 homo sapien
15	23.14	42.5	1022	4	Q9P249	Q9P249 homo sapien
16	22.13	40.6	418	4	Q9G1U9	Q9G1U9 homo sapien

Qabx98 mus musculus  
Qabx8 hirc sapien  
Qabf7 homo sapien  
Qabf6 mus musculus  
Qabf38 homo sapien  
Qabf16 homo sapien  
Qabf99 homo sapien  
Qabf43 mus musculus  
Qabf44 rattus norv  
Qabf52 drosophila  
Qabf53 drosophila  
Qabf54 drosophila  
Qabf99 xeropus lae  
Qabx8 mus musculus  
Qabf7 homo sapien  
Qabf6 mus musculus  
Qabf34 mus musculus  
Qabf7 drosophila  
Qabf28 homo sapien  
Qabf1 drosophila  
Qabf0 drosophila  
Qabx0 gallus gall  
Qabx23 mus musculus  
Qabf1 mus musculus  
Qabf1 mus musculus  
Qabf78 drosophila  
Qabf84 homo sapien  
Qabf99 drosophila  
Qabf84 homo sapien  
Qabf99 hirc sapien  
Qabf98 hirc sapien  
Qabf99 homo sapien  
Qabf99 drosophila  
Qabf99 drosophila

## ALIGNMENTS

## RESULTS

```

ID      Q9H2E6      PRESIMINARY;      PR7: 1030 AA.
AC
CT Q9H2E6;
DT CL-VAR-2001 (TEMBRel. 16, Created)
DT CL-VAR-2001 (TEMBRel. 16, Last sequence update)
DT CL-VAR-2003 (TEMBRel. 23, Last annotation update)
DE Semaphorin SEMA5Aa.
CC Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
CC Q95. TaxID=3606;
RY
..
SEQUENCE FROM N.A.
RP MEDLINE=20564339; PubMed=10393894;
RX Klostermann A., Lutz B., Gertler F., Bell C.;
RT The orthologous human and murine semaphorin 5A-1 p
RT /sema5a-1 bind to the Enabled/Vasodilator-stimula
RT like Protein (EVL) via a novel catalytic-terminal Zyn
RL J. Biol. Chem. 275:39647-39653 (2000).
DR EMB3; AF279656; AAC29178.1;
DR InterPro; IP003659; Plexin-like.
DR InterPro; IP001627; Sema.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR SMART; SM00630; Sema; 1.
DR SEQUENCE 1030 AA. 114369 MW: A57B79C10AEC4B34

```

```
Query Match:      100.0%; Score 5450; DB 4; Length 1030;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

	QY	DB
1	MRSBALLYTLLHFAGAGPEPDESEPSISGHGNYKQYPVFVGHHPGRNTTCQRHLDIQM 60	
1	MRSBALLYTLLHFAGAGPEPDESEPSISGHGNYKQYPVFVGHHPGRNTTCQRHLDIQM 60	
61	IMINMGTYLAARDHITYVDIDTSHTTEEYCCKKTWASROADVDTCKMKGGKHKDECHNF 120	
61	IMINMGTYLAARDHITYVDIDTSHTTEEYCCKKTWASROADVDTCKMKGGKHKDECHNF 120	

y	121	IKVLLKKNDALFVCGTNAENPSCRYKMDLLEPFQDEFSQVACQCYDAKHANVALFADG	130
b	121	IKVLLKKNDALFVCGTNAENPSCRYKMDLLEPFQDEFSQVACQCYDAKHANVALFADG	130
y	181	KLYSATVDFALDAVAYRSLGSSPTLRITKQSKWLEKPYFVQAVDYGDIYVFFREIA	240
b	181	KLYSATVDFALDAVAYRSLGSSPTLRITKQSKWLEKPYFVQAVDYGDIYVFFREIA	240
y	241	VEYNTMGVFFRVAQVCKNDMGSSQVLEKQWTSFLKAPLNCVFGDSHFYFNILQAVT	300
b	241	VEYNTMGVFFRVAQVCKNDMGSSQVLEKQWTSFLKAPLNCVFGDSHFYFNILQAVT	300
y	301	DVIRINGRDVVLATFSTPYKSIPISSAVCAVMDLDAVYTGREKQKPDSTWTFVQDER	360
b	301	DVIRINGRDVVLATFSTPYKSIPISSAVCAVMDLDAVYTGREKQKPDSTWTFVQDER	360
y	361	VPKPRGCCAGSSSLERYATSNRFPDQTLNFKCHPLMDEAVPSIENRPFRLTWVRYEL	420
b	361	VPKPRGCCAGSSSLERYATSNRFPDQTLNFKCHPLMDEAVPSIENRPFRLTWVRYEL	420
y	421	TKIAVDTAAGPYQNHVYVFGSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSEKSYD	480
b	421	TKIAVDTAAGPYQNHVYVFGSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSEKSYD	480
y	481	GVEDKRMGMQDRASSSLYVAFSTCVIKVPLGRCEHKGKCKTCIASRDPYCGWIKEGG	540
b	481	GVEDKRMGMQDRASSSLYVAFSTCVIKVPLGRCEHKGKCKTCIASRDPYCGWIKEGG	540
y	541	ACSHLSNPNRUTFEQDERGNTGCGDCHNSFVALNDSTPLPQNEVSYNTVYGHSSSL	600
b	541	ACSHLSNPNRUTFEQDERGNTGCGDCHNSFVALNDSTPLPQNEVSYNTVYGHSSSL	600
y	594	PSITTSSTAGCYESRGCMQDLWKHLSDSPDTPGLGAVSSHNHQQKKGVIRESYLKQHD	643
b	594	PSITTSSTAGCYESRGCMQDLWKHLSDSPDTPGLGAVSSHNHQQKKGVIRESYLKQHD	643
y	603	PSITTSSTAGCYESRGCMQDLWKHLSDSPDTPGLGAVSSHNHQQKKGVIRESYLKQHD	662
b	603	PSITTSSTAGCYESRGCMQDLWKHLSDSPDTPGLGAVSSHNHQQKKGVIRESYLKQHD	662

RESULT 2  
9P2H9  
D Q9P2H9 PRELIMINARY; PRT: 1049 AA.  
C Q9P2H9  
T 01-OCT-2000 (TEMBLrel. 15, Created)  
T 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
T 01-MAR-2003 (TEMBLrel. 21, Last annotation update)

DE	Hypothetical protein KIAA1368 (Fragment).			
GN	KIAA1368.			
CS	Homo sapiens (Human).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
CC	NCB_TaxID=9606;			
PN	11_1			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=20181126; PubMed=10718198;			
RA	Magase T., Kikuno R., Ishikawa K., Hirosewa M., Ohara C.;			
RT	Prediction of the coding sequences of unidentified human genes.XV.			
RT	The complete sequences of 150 new cDNA clones from brain which code			
RT	for large proteins in vitro.			
RU	DRA Res. 1165-73(2003).			
DR	EMBL: AB017789; BAA92606.1; -			
DR	Genbank: HGNC:10738; SEMA6A.			
DR	InterPro: IPR003659; Plectin-like.			
DR	InterPro: IPR001403; Sema.			
DR	Pfam: PF01403; Sema; 1.			
DR	SMART: SM00423; PSI; 1.			
DR	SMART: SM00630; Sema; 1.			
KW	Hypothetical protein.			
FT	NON_TER 1			
SC	SEQUENCE 1049 AA: 116511; MW: 778120AOCCTA9A8A5A CRC64;			
	Query Match	99.7%;	Score 5431.5;	DB 4; Length 1049;
	Best Local Similarity	98.4%;	Pred. No. 0;	
	Matches 1010;	Conservative 0;	Mismatches 0;	Indels 17; Gaps 1;
OY	1	MPSEALLLYPTLLFPAGAGPEESEPIS:SHGNTYKQYVPVGHKPGPNTTCRRLDIOM	60	
Ob	3	KRSEALLLYPTLLFPAGAGPEESEPIS:SHGNTYKQYVPVGHKPGPNTTCRRLDIOM	62	
OY	61	IVIMNGTILYIARORHIVTVDICTSHTEELVYSGSKLTWKSPQADYDTCYKGRKDFCHNF	120	
Ob	63	IVIMNGTILYIARORHIVTVDICTSHTEELVYSGSKLTWKSRQADYDTCYKGRKDFCHNF	122	
OY	121	KVLLKKNDALFVCGTNAENPSCRYKMDLLEPFQDEFSQVACQCYDAKHANVALFADG	180	
Ob	123	KVLLKKNDALFVCGTNAENPSCRYKMDLLEPFQDEFSQVACQCYDAKHANVALFADG	182	
OY	181	KLYSATVDFALDAVAYRSLGSSPTLRITVKHDSKWLKEEYFVQAVDYGDIYVFFREIA	240	
Ob	183	KLYSATVDFALDAVAYRSLGSSPTLRITVKHDSKWLKEEYFVQAVDYGDIYVFFREIA	242	
OY	241	VEYNTMGKVFPRVAQVCKNDMGSSQVLEKQWTSFLKARLNCSPVGGSHFYFN:LQAVT	300	
Ob	243	VEYNTMGKVFPRVAQVCKNDMGSSQVLEKQWTSFLKARLNCSPVGGSHFYFN:LQAVT	302	
OY	301	DVIRINGRDVVLATFSTPYKSIPIGSAVCAVMDLDAVFTGRFKQSKSPDSTWTFVQDER	360	
Ob	303	DVIRINGRDVVLATFSTPYKSIPIGSAVCAVMDLDAVFTGRFKQSKSPDSTWTFVQDER	362	
OY	361	VPKPRGCCAGSSSLERYATSNRFPDQTLNFKTHPLMDEAVPSIENRPFRLTWVRYEL	420	
Ob	363	VPKPRGCCAGSSSLERYATSNRFPDQTLNFKTHPLMDEAVPSIENRPFRLTWVRYEL	422	
OY	421	TKIAVDTAAGPYQNHVYVFGSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSEKSYD	480	
Ob	423	TKIAVDTAAGPYQNHVYVFGSEKGIILKFLARIGNSGFLNDSLFLEMSVYNSEKSYD	482	
OY	481	GVEDKRMGMQDRASSSLYVAFSTCVIKVPLGRCEHKGKCKKTCIASRDPYCGWKKEGG	540	
Ob	483	GVEDKRMGMQDRASSSLYVAFSTCVIKVPLGRCEHKGKCKKTCIASRDPYCGWKKEGG	542	
OY	541	ACSHLSNPNRSLTFEQDIERGNTGCGDCHNSFVALN-----GHSSSL	563	
Ob	543	ACSHLSNPNRSLTFEQDIERGNTGCGDCHNSFVALNDSPLPDNEMSYNTVYGHSSSL	602	
OY	594	PSITTSSTAGCYESRGGMQDLWKHLSDSPDTPGLGAVSSHNHQQKKGVIRESYLKQHD	643	
Ob	603	PSITTSSTAGCYESRGGMQDLWKHLSDSPDTPGLGAVSSHNHQQKKGVIRESYLKQHD	662	

```

644 QVVFVTLIAIANTLAPVMAVPSGCTVYVCCHERKDVAVVQKKEKELTHSRGMSAVT 703
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
663 QLVPPVTLIAIANTLAPVMAVPSGCTVYVCCHERKDVAVVQKKEKELTHSRGMSAVT 722
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
704 KLSGAFGDTQSCPKPEA---TLPLNNGKLTATGNTAKMLIKADQHLTLTALPTPESTPT 763
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
723 KLSGAFGDTQSCPKPEA---TLPLNNGKLTATGNTAKMLIKADQHLTLTALPTPESTPT 782
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
764 LQKKRKSRSSEWERNOKLEINACTKCMPPKSPVPTDLPPLASFSHIPSPVYVLPITQQ 823
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
783 LQKKRKSRSSEWERNOKLEINACTKCMPPKSPVPTDLPPLASFSHIPSPVYVLPITQQ 842
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
924 GYQHEVYDFKMSVACMALEQCAATLEYKTKKEHLSKSPHNGVJLVNLSLPKVPQ 863
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
943 GYQHEVYDFKMSVACMALEQCAATLEYKTKKEHLSKSPHNGVJLVNLSLPKVPQ 902
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
864 REASLPPGASLSOTLSKLELVHSSSYGVYKRSYFNSLTSRSHOATLKRNTNSN 943
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
903 REASLPPGASLSOTLSKLELVHSSSYGVYKRSYFNSLTSRSHOATLKRNTNSN 962
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
944 SSSLSENQSGFGNDPPAPQVDSQVHSSQSPSGOAVTVSRQPSLNAVNSLTSRSLKET 1003
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
963 SSSLSENQSGFGNDPPAPQVDSQVHSSQSPSGOAVTVSRQPSLNAVNSLTSRSLKET 1022
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1004 PSLKPDVPEKPSAPLSTSMKPNDACT 1030
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1023 PSLKPDVPEKPSAPLSTSMKPNDACT 1049
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 3

```

95Q7: PRELIMINARY: PRT: 1005 AA.
C SE071:
01-MAR-2001 (TREMBLrel. 16, Created)
T 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
T 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
E Axon guidance signal SEMA6A.
N SEMA6A.
S Mus musculus (Mouse).
S Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
V NCBI_TaxID:10090;
V [1]_TaxID:10090;
P SEQUENCE FROM N.A.
P CDSUE=Brain;
X MEDLINE=20364339; PubMed=10933894;
A Klosterrmann A., Lutz B., Gestblom P., Pahl G.
T "The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-
T 1/SEMA6A-1) bind to the Enabled/Nasodulator-stimulated Phosphotyrosine-
T like Protein (EVL) via a novel carboxy-terminal Cxxin-like domain."
T J. Biol. Chem. 275:39547-39553(2000).
T EVBL; AF288666; AAG29494.1; -.
R MGD; MGI:1203727; SemA6a.
R InterPro: IPR001659; Plexin-like.
R InterPro: IPR001627; Sema.
R Pfam: PF01403; Sema; 1.
R SMART: SM00423; PSI; 1.
R SMART: SM00630; Sema; 1.
Q SEQUENCE 1005 AA, 111758 MW, 575699279458790 CRR64;

```

```

Query Match 92.2%; Score 5024.5; DB 1; Length 1005;
Best Local Similarity 91.9%; Pval: No. 0;
Matches 948; Conservative 26; Mismatches 30; Indels 27; Gaps 2;

```

```

1 MRSEALLVFTLLHFAGAGFPDSEPSISHGNYTKQYVFGVGHKPGRNTTORHLDIQ: 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1 MRPAALLCLTLHCAGAGFPDSEPSISHGNYTKQYVFGVGHKPGRNTTORHLDIQ: 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 IMIMNCTLYAARDHIYTVDDITSHFEIYVCSKKLTWKSQADVDTCRMKGKHKECHNF 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 IMIMNCTLYAARDHIYTVDDITSHFEIYVCSKKLTWKSQADVDTCRMKGKHKECHNF 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

121 LKWLKKNDALFVCGTNAPNBSGKVKYDTLSPGDFESNARCPDYDAKHANVALEADG 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121 LKWLKKNDALFVCGTNAPNBSGKVKYDTLSPGDFESNARCPDYDAKHANVALEADG 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
191 KLYSATVTDPLAIDAVYRSLGESPTLRTVKHDSKWLKEFYFVCAVDYGYIYFFFEREIA 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
191 KLYSATVTDPLAIDAVYRSLGESPTLRTVKHDSKWLKEFYFVCAVDYGYIYFFFEREIA 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
241 VEYNTMKVYFPRVAQVCKNDMGSGQVLEKQWTSFLKAPLANSVPSESHFYFNILQAVT 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
241 VEYNTMKVYFPRVAQVCKNDMGSGQVLEKQWTSFLKAPLANSVPSESHFYFNILQAVT 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 DVIRINGRQVYLFATESFYNSIPGSAYCAVDMLDIASVFTGRFKECKSPDSTWTPVDER 360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 DVIRINGRQVYLFATESFYNSIPGSAYCAVDMLDIASVFTGRFKECKSPDSTWTPVDER 360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
361 VPKPRPCCAGSSSLERYATSNFFPDOTLNFIKTHPLMDEAVPSIENRPFELRTMYRYRL 420
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
361 VPKPRPCCAGSSSLERYATSNFFPDOTLNFIKTHPLMDEAVPSIENRPFELRTMYRYRL 420
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
421 TKZAVDTAAGPYQNHITVYFLGSEKGIILKELARIGNSGFLNLSLFEEMSVYNSKCSYD 480
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
421 TKZAVDTAAGPYQNHITVYFLGSEKGIILKELARIGNSGFLNLSLFEEMSVYNSKCSYD 480
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
481 GVEKRIWGMQLODRASSSLYVAFSTCVIKVPLGRCEPHGKCKTCTIASRDPYCGWKEGG 540
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
481 GVEKRIWGMQLODRASSSLYVAFSTCVIKVPLGRCEPHGKCKTCTIASRDPYCGWKEGG 540
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
541 ACSHLSPNSRLTTECDIERGNTDGLGCHNSFVALNGHSSSLPSTTSSDAGEYER 600
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
541 ACSHLSPNSRLTTECDIERGNTDGLGCHNSFVALNGHSSSLPSTTSSDAGEYER 600
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
545 SCAHLSPLSALTTECDIERGNTDGLGCHNSFVALNGHSSSLPSTTSSDAGEYER 574
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
501 GGWLKWHLLDSPDSTPLGAVSHNHQCKKGVIRESYLKGHDLQVPTLLA-AVILAFV 660
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
575 GGWLKWHLLDSPDSTPLGAVSHNHQCKKGVIRESYLKGHDLQVPTLLA-AVILAFV 634
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
561 MGAVFSGTIVYVCDHREKDVAVVQKKEKELTHSRGMSAVTVLXSLGDTQSKDKPE 720
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
561 MGAVFSGTIVYVCDHREKDVAVVQKKEKELTHSRGMSAVTVLXSLGDTQSKDKPE 720
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
595 MGAVFSGTIVYVCDHREKDVAVVQKKEKELTHSRGMSAVTVLXSLGDTQSKDKPE 694
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
721 AILPLMHNGKLTATGNTAKMLIKADQHLTLTALPTPESTPTLQKKRKPGRSREARN 780
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
721 AILPLMHNGKLTATGNTAKMLIKADQHLTLTALPTPESTPTLQKKRKPGRSREARN 780
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
781 QNLNACTKMPKSPVPTDLPPLASFSHIPSPVPTDLPPLASFSHIPSPVPTDLPPLASFS 839
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
781 QNLNACTKMPKSPVPTDLPPLASFSHIPSPVPTDLPPLASFSHIPSPVPTDLPPLASFS 839
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
755 QNLNACTKMPKSPVPTDLPPLASFSHIPSPVPTDLPPLASFSHIPSPVPTDLPPLASFS 814
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
840 QMALEDQAATLEYKTKKEHLSKSPHNGVJLVNLSLPPKVPQREASLPPGASLSOTG 899
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
840 QMALEDQAATLEYKTKKEHLSKSPHNGVJLVNLSLPPKVPQREASLPPGASLSOTG 899
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
815 QMALEDQAATLEYKTKKEHLSKSPHNGVJLVNLSLPPKVPQREASLPPGASLSOTG 874
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
900 LSKRLEMHSSSYGVYKRSYFNSLTSRSHOATLKRNTNSNSSLKRNQSGFGNDP 959
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
875 LSKRLEMHSSSYGVYKRSYFNSLTSRSHOATLKRNTNSNSSLKRNQSGFGNDP 934
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
960 PPAPQVDSQVHSSQSPSGOAVTVSRQPSLNAVNSLTSRSLKETPSLKPDPVPRKPSAPL 1019
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
915 PPAPQVDSQVHSSQSPSGOAVTVSRQPSLNAVNSLTSRSLKETPSLKPDPVPRKPSAPL 994
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1020 STSMKPNDACT 1030
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
995 STSMKPNDACT 1005
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 4  
Q96SK4  
ID Q96SK4 PRELIMINARY: PRT: 699 AA.  
AC Q96SK4:  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Hypothetical protein FJ24595.



```

482 NTNSNSSILSKNSGEGDNEPAPQRCVSTQVSSSSGAGVAVSRQPSLNAVNSJTR 541
998 SGLKTPSKPQVPPKPSFAPJSTSMKPNDACT 1030
542 SGLKTPSKPQVPPKPSFAPJSTSMKPNDACT 574

RESULT 6
88BUC0
D C8BUC0 PRELIMINARY; PRT; 587 AA.
C C9BUC0;
T 01-MAR-2003 (TREMBlrel. 23, Created);
T 01-MAR-2003 (TREMBlrel. 23, Last sequence update);
T 01-MAR-2003 (TREMBlrel. 23, Last annotation update);
E Sema domain.
S Mus musculus (Mouse).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
X NCBI_TaxID=10330;
W
P SEQUENCE FROM N.A.
P STRAIN=C57BL/6J; TISSUE=Cerebellum;
C MEDLINE=22354683; PubMed=12466851;
A The FANTOM Consortium.
A "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
R Nature 420:563-573(2002);
R EMBL; AK084711; BAC38582.17..
Q SEQUENCE 587 AA; 66044 MW; 8E9C7B102C4D897A CRC64;

Query Match 54.4%; Score 2963; DB 11; Length 587;
Best Local Similarity 95.1%; Pred. No. 1,8e-232;
Matches 549; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

1 MRSEALLLYTLHFAGAGFPDESEPISSHGNYTKQYVVFVGHKPGRNNTQRRHLDIQ 60
1 MPAALLLCCTLLHCAGAGFPDESEPISSHGNYTKQYVVFVGHKPGRNNTQRRHLDIQ 60
61 IMXNGTLVIAARCHIYTVDTSTHTTEIYCSKKLTWKSQADVDTCRMKGKDECHNF 120
61 IMXNRTLVIAARCHIYTVDTSTHTTEIYCSKKLTWKSQADVDTCRMKGKDECHNF 120
121 IKVLLKNDALFVCGTNAPNSCRNYKMDTLEPFGDEFSGMARCPYDAKHANVAFAG 180
121 IKVLLKNDCTLFVCGTNAPNSCRNYKMDTLEPFGDEFSGMARCPYDAKHANVAFAG 180
181 KLYSATVTDFLAIDAVIYRSLGSPDLRTVYKHSKWLKEPYFYQAVDYGDIYFFPREIA 240
181 KLYSATVTDFLAIDAVIYRSLGSPDLRTVYKHSKWLKEPYFYQAVDYGDIYFFPREIA 240
241 VEYNTMGKVVFPRAVQVCKNDMGSGORVLEKQWTSFLKARLNCSPGDSHFYFNILCAVT 300
241 VEYNTMGKVVFPRAVQVCKNDMGSGORVLEKQWTSFLKARLNCSPGDSHFYFNILCAVT 300
301 DVIRINGRDVVLATSTPYNSIPGSAVCAYDMLEDIASVFTGRFKEQKSPDSTWTPVEDR 360
241 VEYNTMGKVVFPRAVQVCKNDMGSGORVLEKQWTSFLKARLNCSPGDSHFYFNILCAVT 300
301 DVIRINGRDVVLATSTPYNSIPGSAVCAYDMLEDIASVFTGRFKEQKSPDSTWTPVEDR 360
361 VPKPRPGCCAGSSSLEKYATSNFPDDTLNFITKTHPLMDEAVPSINRPWFLRTVYVRL 420
361 VPKPRPGCCAGSSSLEKYATSNFPDDTLNFITKTHPLMDEAVPSINRPWFLRTVYVRL 420
421 TKIADVTAAGPYQNHVTVPLGSEKGIILKFLARIGNSGFLNDSLFLSESVVYNSEKCSYD 480
421 TKIADVNAAGPYQNHVTVPLGSEKGIILKFLARIGNSGFLNDSLFLSESVVYNSEKCSYD 480
481 GVEDKRIINGVQLDRASSSLYVAFSTCVIKVPLGRCHRGKCKKTCIASRDPYCGWKEGG 540
491 GVEDKRIINGVQLDRASSSLYVAFSTCVIKVPLGRCHRGKCKKTCIASRDPYCGWKEGG 540
541 ACSHLSPNRLTFEQDIERGNTDGLGCHNSFVALNDISTP; 582
```

```

541 SCALSPSLRLTFEQDIERGNTDGLGCHNSFVALNG 577
PRELIMINARY; PRT; 605 AA.
C C8EXZ7;
C C9EXZ7;
T 01-MAR-2003 (TREMBlrel. 23, Created);
T 01-MAR-2003 (TREMBlrel. 23, Last sequence update);
T 01-MAR-2003 (TREMBlrel. 23, Last annotation update);
E Sema domain.
C Mus musculus (Mouse).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
X NCBI_TaxID=10090;
W
P SEQUENCE FROM N.A.
P STRAIN=C57BL/6J; TISSUE=Cerebellum;
C MEDLINE=22354683; PubMed=12466851;
A The FANTOM Consortium.
A "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
R Nature 420:563-573(2002);
R EMBL; AK042751; BAC31351.1..
Q SEQUENCE 605 AA; 62299 MW; B4FE713BE8AF6E90 CRC64;

Query Match 54.3%; Score 2961; DB 11; Length 605;
Best Local Similarity 94.5%; Pred. No. 2,8e-232;
Matches 550; Conservative 14; Mismatches 18; Indels 3; Gaps 0;

1 MRSEALLLYTLHFAGAGFPDESEPISSHGNYTKQYVVFVGHKPGRNNTQRRHLDIQ 60
1 MRPAALLLCCTLLHCAGAGFPDESEPISSHGNYTKQYVVFVGHKPGRNNTQRRHLDIQ 60
61 IMXNGTLVIAARCHIYTVDTSTHTTEIYCSKKLTWKSQADVDTCRMKGKDECHNF 120
61 IMXNRTLVIAARCHIYTVDTSTHTTEIYCSKKLTWKSQADVDTCRMKGKDECHNF 120
121 IKVLLKNDALFVCGTNAPNSCRNYKMDTLEPFGDEFSGMARCPYDAKHANVAFAG 180
121 IKVLLKNDCTLFVCGTNAPNSCRNYKMDTLEPFGDEFSGMARCPYDAKHANVAFAG 180
181 KLYSATVTDFLAIDAVIYRSLGSPDLRTVYKHSKWLKEPYFYQAVDYGDIYFFPREIA 240
181 KLYSATVTDFLAIDAVIYRSLGSPDLRTVYKHSKWLKEPYFYQAVDYGDIYFFPREIA 240
241 VEYNTMGKVVFPRAVQVCKNDMGSGORVLEKQWTSFLKARLNCSPGDSHFYFNILCAVT 300
241 VEYNTMGKVVFPRAVQVCKNDMGSGORVLEKQWTSFLKARLNCSPGDSHFYFNILCAVT 300
301 DVIRINGRDVVLATSTPYNSIPGSAVCAYDMLEDIASVFTGRFKEQKSPDSTWTPVEDR 360
301 DVIRINGRDVVLATSTPYNSIPGSAVCAYDMLEDIASVFTGRFKEQKSPDSTWTPVEDR 360
361 VPKPRPGCCAGSSSLEKYATSNFPDDTLNFITKTHPLMDEAVPSINRPWFLRTVYVRL 420
361 VPKPRPGCCAGSSSLEKYATSNFPDDTLNFITKTHPLMDEAVPSINRPWFLRTVYVRL 420
421 TKIADVTAAGPYQNHVTVPLGSEKGIILKFLARIGNSGFLNDSLFLSESVVYNSEKCSYD 480
421 TKIADVNAAGPYQNHVTVPLGSEKGIILKFLARIGNSGFLNDSLFLSESVVYNSEKCSYD 480
481 GVEDKRIINGVQLDRASSSLYVAFSTCVIKVPLGRCHRGKCKKTCIASRDPYCGWKEGG 540
541 ACSHLSPNRLTFEQDIERGNTDGLGCHNSFVALNGHSSSL 582
541 SCALSPSLRLTFEQDIERGNTDGLGCHNSFVALNDISTP; 582
```

```

RESULT 8
96SY4
D Q96SY4 PRELIMINARY: PRT: 562 AA.
C Q96SY4:
T 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
T 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
T 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
S Hypothetical protein FJ414565.
E Homo sapiens (Human).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
X NCBI_TaxID=9606;
Y .
F SEQUENCE FROM N.A.
A Isogai T., Ota T., Nishikawa T., Sugiyama T., Otsuki T., Suzuki Y.,
A Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
A Watanabe S., Hosobuchi T., Kaku Y., Kodaira K., Kondo H., Sugawara M.,
A Takahashi M., Chiba Y., Ishida S., Yuzakawa K., Ono Y., Takiguchi S.,
A Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
A Yamamoto C., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
A Kinomiyama K., Iwayanagi T.,
A "NEDO human cDNA sequencing project.",
I Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
R EMBL: AK024713; BA355136.1;
R InterPro: IPR003639; Flexin-like.
R SMART: SMC423; PSI; ..
R Hypothetical protein.
W SEQUENCE 562 AA; 61313 MW; 6AB3695FAD1DD78A CR664;

Query Match 54.2%; Score 2952; DB 4; Length 562;
Best Local Similarity 102.0%; Pred. No. 136-231;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

469 MSYNEKSYGVDEKRIKNGVQLDRASSSLVAFSTGVKPLGRCERHGKCKTCIAS 529
: MSYNEKSYGVDEKRIKNGVQLDRASSSLVAFSTGVKPLGRCERHGKCKTCIAS 60
529 RDPYCKWIKEGGACSHLSNRLTFEQD-ERGNLTDLGDCRHSFVALNCHSSLSLSTTT 589
: RDPYCKWIKEGGACSHLSNRLTFEQD-ERGNLTDLGDCRHSFVALNCHSSLSLSTTT 120
61 RDPYCKWIKEGGACSHLSNRLTFEQD-ERGNLTDLGDCRHSFVALNCHSSLSLSTTT 120
589 SDSTAQEGYESRGKGLDWKHLDSPTDPLGAVSHHCKKGVIRESYLKCHDQLVPV 649
121 SDSTAQEGYESRGKGLDWKHLDSPTDPLGAVSHHCKKGVIRESYLKCHDQLVPV 180
649 TLLAIAVILAFVNGAVFSGITVYVCDHRKGVAVVQVCKEKLTHSRGSMSSVTKLSG 709
181 TLLAIAVILAFVNGAVFSGITVYVCDHRKGVAVVQVCKEKLTHSRGSMSSVTKLSG 240
709 FGDTSQKPKFEAILTFLMNGKLTATGNTAKVLIKADQHHLDLTALPTPESTPTLQCKR 769
241 FGDTSQKPKFEAILTFLMNGKLTATGNTAKVLIKADQHHLDLTALPTPESTPTLQCKR 300
769 KPSRGSRERNONLNACTKDMPPKSGPVIPDPLPRASPSHIPSVVLPITQCGYCHE 829
301 KPSRGSRERNONLNACTKDMPPKSGPVIPDPLPRASPSHIPSVVLPITQCGYCHE 360
829 YVDPQKXSEVAQMALEDQAATLEYKTKHELSKSPNHNGLVENLDSLPKVPQREASL 889
361 YVDPQKXSEVAQMALEDQAATLEYKTKHELSKSPNHNGLVENLDSLPKVPQREASL 420
889 GPPGASLSQGLSKRLMHSSSYGVYKSYPTNSLTTRSHQATTLKRNNTNSSNSHLS 948
421 GPPGASLSQGLSKRLMHSSSYGVYKSYPTNSLTTRSHQATTLKRNNTNSSNSHLS 480
949 RNSGFRGNGNPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTFSLKP 1009
481 RNSGFRGNGNPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTFSLKP 540
1009 DVPPKPSFAPLSTSMKPNDACT 1030
541 DVPPKPSFAPLSTSMKPNDACT 562

RESULT 10
Q96SY4
D Q96SY4 PRELIMINARY: PRT: 562 AA.
C Q96SY4:
T 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
T 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
T 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
S Hypothetical protein FJ414565.
E Homo sapiens (Human).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
X NCBI_TaxID=9606;
Y .
F SEQUENCE FROM N.A.
A Isogai T., Ota T., Nishikawa T., Sugiyama T., Otsuki T., Suzuki Y.,
A Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
A Watanabe S., Hosobuchi T., Kaku Y., Kodaira K., Kondo H., Sugawara M.,
A Takahashi M., Chiba Y., Ishida S., Yuzakawa K., Ono Y., Takiguchi S.,
A Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
A Yamamoto C., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,
A Kinomiyama K., Iwayanagi T.,
A "NEDO human cDNA sequencing project.",
I Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
R EMBL: AK024713; BA355136.1;
R InterPro: IPR003639; Flexin-like.
R SMART: SMC423; PSI; ..
R Hypothetical protein.
W SEQUENCE 562 AA; 61313 MW; 6AB3695FAD1DD78A CR664;

Query Match 54.2%; Score 2952; DB 4; Length 562;
Best Local Similarity 102.0%; Pred. No. 136-231;
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

469 MSYNEKSYGVDEKRIKNGVQLDRASSSLVAFSTGVKPLGRCERHGKCKTCIAS 529
: MSYNEKSYGVDEKRIKNGVQLDRASSSLVAFSTGVKPLGRCERHGKCKTCIAS 60
529 RDPYCKWIKEGGACSHLSNRLTFEQD-ERGNLTDLGDCRHSFVALNCHSSLSLSTTT 589
: RDPYCKWIKEGGACSHLSNRLTFEQD-ERGNLTDLGDCRHSFVALNCHSSLSLSTTT 120
61 RDPYCKWIKEGGACSHLSNRLTFEQD-ERGNLTDLGDCRHSFVALNCHSSLSLSTTT 120
589 SDSTAQEGYESRGKGLDWKHLDSPTDPLGAVSHHCKKGVIRESYLKCHDQLVPV 649
121 SDSTAQEGYESRGKGLDWKHLDSPTDPLGAVSHHCKKGVIRESYLKCHDQLVPV 180
649 TLLAIAVILAFVNGAVFSGITVYVCDHRKGVAVVQVCKEKLTHSRGSMSSVTKLSG 709
181 TLLAIAVILAFVNGAVFSGITVYVCDHRKGVAVVQVCKEKLTHSRGSMSSVTKLSG 240
709 FGDTSQKPKFEAILTFLMNGKLTATGNTAKVLIKADQHHLDLTALPTPESTPTLQCKR 769
241 FGDTSQKPKFEAILTFLMNGKLTATGNTAKVLIKADQHHLDLTALPTPESTPTLQCKR 300
769 KPSRGSRERNONLNACTKDMPPKSGPVIPDPLPRASPSHIPSVVLPITQCGYCHE 829
301 KPSRGSRERNONLNACTKDMPPKSGPVIPDPLPRASPSHIPSVVLPITQCGYCHE 360
829 YVDPQKXSEVAQMALEDQAATLEYKTKHELSKSPNHNGLVENLDSLPKVPQREASL 889
361 YVDPQKXSEVAQMALEDQAATLEYKTKHELSKSPNHNGLVENLDSLPKVPQREASL 420
889 GPPGASLSQGLSKRLMHSSSYGVYKSYPTNSLTTRSHQATTLKRNNTNSSNSHLS 948
421 GPPGASLSQGLSKRLMHSSSYGVYKSYPTNSLTTRSHQATTLKRNNTNSSNSHLS 480
949 RNSGFRGNGNPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTFSLKP 1009
481 RNSGFRGNGNPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTFSLKP 540
1009 DVPPKPSFAPLSTSMKPNDACT 1030
541 DVPPKPSFAPLSTSMKPNDACT 562
```

```
Q96T04 PRELIMINARY; PRT: 507 AA.
Q96T04
01-DEC-2001 (TrEMBLrel. 19, Created:
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 21, Last annotation update)
E Hypothetical protein F3J4533.
S Homo sapiens (human).
C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
X NCBI_TaxID=9606;
P SEQUENCE FROM N.A.
A Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuka T., Suzuki Y.,
A Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
A Watanabe S., Hosokawa M., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
A Takahashi Y., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
A Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
A Yamamoto C., Wakamatsu A., Nakamura Y., Nagahara K., Masuko Y.,
A Kinomiyama K., Iwayanagi T.,
I "NEDO human cDNA sequencing project.";
J Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
R EMBL; AK027439; BAB55111.1;
R InterPro; IPR003659; Plectin-like.
R SMART; SM00423; PS: 1.
M Hypothetical protein.
Q SEQUENCE 507 AA; 55466 MW; 8CC567B438C51B39 CRC64;

Query Match 48.1%; Score 2621.5; DB 4; Length 507;
Best Local Similarity 90.2%; Pred. NO. 9.9e-205;
Matches 507; Conservative C; Mismatches C; Indels 55; Gaps 1;

469 XSVYNSEKCSYDGVQEKRMGMQLDASSLYAFSTCVKPLGCRGKCKTCTIAS 528
1 XSVYNSEKCSYDGVQEKRMGMQLDASSLYAFSTCVKPLGCRGKCKTCTIAS 60
529 RDPYCGMKEGGACSHLSPNSRIFPDQIERGNTGLGCHGSPVALNGHSSSLPSTTT 588
61 RDPYCGMKEGGACSHLSPNSRIFPDQIERGNTGLGCHGSPVALNGHSSSLPSTTT 588
589 SDSTAQEGYSRGMGLDWKHLDEPSTDFLGAVSHNHOLKGVPESEYKQHQCLQPV 648
109 -----GVRESYKGRHQCLQPV 125
649 TLLAIAVILAFVNGAVFSGITVYCVDRPKQVAVYCRKEKELTHSPGSSYKSLSL 708
126 TLLAIAVILAFVNGAVFSGITVYCVDRPKQVAVYCRKEKELTHSPGSSYKSLSL 135
709 FGDQSKDPKRAILTLFLMHGKGLATGNTAKNLKADQHLGLTALPTFESTPLQQR 768
126 FGDQSKDPKRAILTLFLMHGKGLATGNTAKNLKADQHLGLTALPTFESTPLQQR 745
769 KPSRGSSEKERNQGLNACTKDYIEMGSEVPTDLPIKASESHPEVWVLPITQCYQHE 828
246 KPSRGSSEKERNQGLNACTKDYIEMGSEVPTDLPIKASESHPEVWVLPITQCYQHE 305
829 YVQPKMSEVAQMALEDCAATLKYTKIKELHLSKSPNHNGLVENVLDSIFPKVFGREASL 886
306 YVQPKMSEVAQMALEDCAATLKYTKIKELHLSKSPNHNGLVENVLDSIFPKVFGREASL 365
889 GPPGASLCTGLSKRLMHSSSYGYCYKASYPTNLSLTHSHQATLTKRNTNTHSSSHLS 948
366 GPPGASLCTGLSKRLMHSSSYGYCYKASYPTNLSLTHSHQATLTKRNTNTHSSSHLS 425
949 RNQSGFGSDPPAPQVDSIQVESSQSGQAVTVSQSLKAVNSLTPRSGKRTPLSKP 1009
426 RNQSGFGSDPPAPQVDSIQVESSQSGQAVTVSQSLKAVNSLTPRSGKRTPLSKP 465
1009 DVPPKPSFAFLSTSMKENDACT 1030
486 DVPPKPSFAFLSTSMKENDACT 507
```

RESULT 11

```
Q8NFY4 PRELIMINARY; PRT: 1073 AA.
Q8NFY4
01-OCT-2002 (TrEMBLrel. 22, Created:
01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Semaphorin 6D isoform 4.
GN SEMA6B.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [L:1073]
RF SEQUENCE FROM N.A.
RC [L:1073]
RA CC X., Zhai Y., Wei H., Yu Y., Tang F., He F.;
RT "Homo sapiens semaphorin 6D isoform 4 (SEMA6B.4) mRNA, complete cds.";
BL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
EP EMBL; AF389429; AAM69452.1;
DR InterPro; IPR001627; Sema.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00630; Sema; 1.
SQ SEQUENCE 1073 AA; 119672 MW; 7DCE4DFC5BF7CF9E CRC64;

Query Match 43.4%; Score 2368; DB 4; Length 1073;
Best Local Similarity 46.0%; Pred. NO. 1.5e-183;
Matches 512; Conservative 159; Mismatches 368; Indels 134; Gaps 30;

CY 1 YRSEALLLYFTLL---HFAGAGPEPSEPTISISHGNYTKQYVFGHKPGRNITQHRRLD 57
DS 1 XAVPELLCAYILLNNVCLRAVSPEDDEPLNTVDYHSQYPVFRG-RPSGNEQ-RLRLD 58
CY 58 IQMIMNMGLTYIARDRHYTVDTSHTBEIYCSKKLTWKSQAQVDTCRMKGKHKHDEC 117
DS 59 FQLMKIRDTLYIAGRDQVYTNLNMKPTETVFNKLTWRSQQRENCAMKGRHKHDEC 118
CY 118 HNFYKVLKKNDDALFVCGTNAPNPNCRVMDTLEPFGDFSGMARCPYDAKHANVLF 177
DS 119 HNFYKVPNDENWVFCGTAAPNPNCRVYRSLTLEYDGEESGLARCFDARQTNVLF 178
CY 178 ADGKLYSATVTDLAIDAVIYRSLGESPTLRTVYKDSKWLKSPYFVQAVDYGEYVFFPR 237
DS 179 ADGKLYSATVADFLASDAVYRSMGPGSALRTIKYDSKWLKSPYFVQAVDYGEYVFFPR 236
CY 238 EIAVENTMKGVFPVPAQVCKNDXGSGORVLEKQWTSFLKALNCSVPGDSHFENIQ 297
DS 239 EIAVEHNKQKATSVARICRKNXGSGORVLEKQWTSFLKALNCSVPGDSHFENIQ 298
CY 298 AVTDVIRINGRDVVLATFSTFYNSIPGSACVAYMDIASVFTGFEKQKSPDSTWTPV 357
DS 299 SITTEIQNGIPTVGVPTTCLNSIFGSACVAFSMDDEKVPKGRFKEQKTPDSVWTAVP 358
CY 358 DERVPKPRFCCAGSSSLERYATNSFPDGLNFIKTHLMDEAVPSIENRWFLEMTVP 417
DS 359 ECKVPKPRFCCAGKAGHIAEAYKTSIDFPDETSLFKSHPLMDSAVPPIADEPMFTKTRV 418
CY 418 YSLTKIAVDTAAGPYQNHVTVVUGSEKGLIKFLARIGNSGF-LNDSLPLEFMSVYNSK 476
DS 419 YSLTKIAVDSASAPYQNYTVIFVSGSAGVNLKVLAK--TSPFSLNDSVLEIEAINHAK 476
CY 477 CSYDGVDEKRMGMQLDASSLYVAFSTCVKPLGCRGKCKTCTASRDYPCYGI 536
DS 477 CSAENEEDKKVSLQ-DKDIHIALYVAFSSCIIRIPLSRCERYGSKKCIASRDCYCGWL 536
CY 537 KEGACSHLSPNSRIFTQDIERGNTGLGCHGSPVALNGHSSSLPSTTTSSTAGE 595
DS 537 SQ-GSGGRVTFGLAEQYEQTEFGNTAHLGQCH-----ELPTSTTPD----- 579
CY 596 GYBSRSG-----MCD-MKHL-----DSP-----DSTSP 618
DS 580 -YKIFGPTSCXEVSSSVTTMASIFETPKVIDTWPKJTSRKFVVDNDNTSFTTTP 638
CY 619 LGAVSSHNHCKKGVIRESYLKGHDQLVPVTLIAIAVILAFVNGAVFSGTIVYCVCD-HR 677
```

QY	178	ADGKYSATVTFDA--DAVYVRSJGESPTJRTVKHDSKMLKEPYPVQAVCYGYIYFFER	237
DB	179	ADGKYSATVTFDA--DAVYVRSJGESPTJRTVKHDSKMLKEPYPVQAVCYGYIYFFER	238
QY	238	EIAVEYNTKGVVPRVAQVCKUDMGSGSORVLEKQWTSFLKARLAKCSVPGDSHPFNILQ	297
DB	239	EIAVEHNNCGKAVYSRVAR:CKNDMGSGORVLEKHWTSFLKARLAKCSVPGDSFFVFDVILQ	298
QY	298	AVTVIRINGRDVVVLATFTSPYNSIFGSAYCAVDMLDIASVFTGREKEOKSPDSTWTPVE	357
DB	299	SITD:IOINIGIPVGVFTTCLNSIFGSAYCAFSMODDIEKVFGRKEQKTPDSVWTAIV	358
QY	358	DERVKEPRPOCCAGSSSLERYATSNPEPDTLNFIKTHPELMOEAVPSIFNRPWPLSTWVR	417
DB	359	ECKVKEPRPOCCAGHGLAFAYKTSIDFPDETSLFIKSHPLMDSAVPPADCEWFTKTRVR	418
QY	418	YSL-KIAVDTAAAGPYQKHTVVF-GSKGKIILKFLARIGNSGF--LNDLSLEBMSVNSK	476
DB	419	YRLTAIVSDHSAGPYQNYTVIFVGSAGVLYKJAK--TSPSLNDSVLEBEIAYNHAK	476
QY	477	CSVDGVECKINGQVQDRASSSLYVAFSTCVIKVPLGRGERHGKCKKTCIASRDYCGWI	536
DB	477	CSAENEEDKXVLSLOLCKDHALYVAFSSCIIRIPLSRERYGSKCKSCIASRQVCGWJ	536
QY	537	KEGACSHLSPNSLT-FCQIERGNTDGLGCHNSFVALNGHSSSLSTSTTSSTAGE	595
DB	537	SQ-GSCGRVTFQMLAEGYEQSTFGNTAHLGQCH-----EIGTSTTTPD-----	579
QY	596	GYESRGOMLQWKHLDSFQSTDPGLGAVSHNHQDKKGVIPESLYLKGHGLVPAVT-LAI	655
DB	580	YKIFGG-----PTS-----GVRWEVQSGESNQVMHVLITCV	612
QY	656	LAAPMGAVFSGITVFCVD-HRRKDVAVVQREKEJLTHSRREGSNSSVTLSLFG----	710
DB	613	FAAFVLGAFIAGVAVCYRDPVFKRKI--HKQAESAGCTDSSGSAKJAKGLFQSPYK	670
QY	711	DTQSKDPKEAULTPLMHNGKLA:TPGNTAKXJLKAQOHHLDTALPTEPTPTLOKKRP	770
DB	631	EYQNTDPSKLYSNLTSPKELPKNSETKSMVCHRGQPPPELAALPTPESTPVHOKTQ	730
QY	771	SRGSEWERKONLINAQTRDPMGSPVITDPLFASPSHIPSVVVLPTQOCHGEY-	829
DB	731	AKMSHSEKAHGH--GASRKCTPQFFSSPPHSL--SHGHIPSA-VLNFNATHDYNTSPS	786
QY	830	-----VDOP---KASEVACMALEDAAATLEYKTIKELSSKSPK-----	865
DB	787	NSNAHKAEXKLQNTCHPLTKYSSKGRHRRSVDSRNTJ--NCLLXHLNDSNPKAINMD:	844
QY	856	---HGVLVNLDSL---PKVPQREASLQPPGASLSQTLGSKLEMHSSSY---GVY	916
DB	845	QMAHQNMLDPMGSMSEVPKVPKNREASLYSPSTLPRNSPTKRVQVPTTPGVPMTSLER	904
QY	917	KSYFTNLTASHQATLTKRNNTNNSNSSLHLSNQSGRQD-NPFPAPORVDSIQVHSSC	975
DB	925	QRYHKNSSOR-HSISAVPK-NLNSPQVLLSRQSMNRGGYVPTPGAKVDYIQ-----	957
QY	976	PSGQAVTVSRQPSLNAVNSLT-----RSGLKRTPLSKPDVPEKPSAPLSTSMKP	1025
DB	958	---GTFVSHLQPSLRQSSYTSNGTLPRTGLKRTPLSKPDVPEKPSVVPQPSVAP	1011
RESULT 13			
Q8NFY6			
ID	Q8NFY6	PRELIMINARY;	PRT; 998 AA.
AC	Q8NFY6		
CT	01-OCT-2002	TrEMBLrel. 22, Created!	
CT	01-OCT-2002	TrEMBLrel. 22, Last sequence update!	
CT	01-MAR-2003	TrEMBLrel. 23, Last annotation update!	
DE	Semaphorin 6D isoform 2.		
GN	SEMA6D.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		





```

418 YRUTKIAVDTAGAPYORHNTYVFLSEKGGTILKFAIRIGNSGF-LNDSLFLEMSVYNSEK 476
419 YRUTALASVDSAGPYONYTVTFVSEAGVYLVAK--TSPPSLNDSVLEETEAHNAK 476
477 CSYDGVEDKRIKMGVCLDRASSLYVAFSTVIKVPLOKCRHKKCTIARSDPYCGWI 536
477 CSAENEDKKVLSIQLOKHIALYVAFSSCTIRPLSRCEYVGSCKSCIAERDPYCGWI 536
537 KEGGACSHLSPNSRLT-----FEQDIERGNTDGLGDCHNSFVALNGHSSSL 582
537 SQ-GSCGRVTPGMLLTEDFFAFPHNSAEGYEQDTFGNTAHLCDCRG----- 583
583 LPSTTTSDTAQEGYERGGMLDKHLLDSDPDPJGAVSSHNHQCKGV-RESYLKGH 642
584 -----VRWE--VQSGES----- 593
643 DQGVENVTLALAVILAFMGAVFSGITVYVCD-HRKIVAVYQREKELTHSRGSSVS 701
594 NQWVMVYLITCFAMVVLGAFAGVAVYCYRDMFVKPKI--HKDAESAQSCITSSGS 651
702 VTKLSGLFG---DTQSKDKRKPPEALTFHMHNGKLATEGNTAKYLKADQHH-LDLALPT 757
652 FAKLNGLFDSEVKEVQNIQSPKYSKLLTSRKLPPNGSDTKSMVMHFGQPELAALPT 711
758 PESTPTLQCKRKPGRGREMERNQNLINACTKQVPMGSPVITPDJPLRASPSH:PSVW 817
712 PESTPVLLHCKLQAMKSHSEKAHG--GASRKETPOFFSPSPPHSP--SHGHIPSAIV 778
818 LP-TQCGYQHE?-----VDQP---KQSEVAQMA:EDQAAATLEYVTHREHLS 860
778 LPNATHDYNTSFSNSNAHKAEEKLQNI:DHPLTKGSSKREHRSVDSRNTL--NLLKHLN 836
861 SKSPN-----HGVN:LVENLDSL---PPKVPQREASLGPFGASLSQTGLSKSLEM 906

```

## ESULT 15

```

9P249 PRELIMINARY; PR7; 1022 AA.
C Q9P249;
T C1-CCT-2000 (TrEMBLrel. 15, Created)
T 01-CCT-2001 (TrEMBLrel. 18, Last sequence update)
T 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
E Hypothetical protein KIAA1479 (P-fragment).
N KIAA1479.
N Homo sapiens (Human);
S Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelesstori;
X Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
X NCBI_TaxID=9606;
N [1].
P SEQUENCE FROM N.A.
X MEDLINE=20277482; PubMed=10919333;
X Nagase T., Kikuno R., Ishikawa K., Hirose M., Chava G.:
X "Prediction of the coding sequences of unidentified human
X genes. XVII. The complete sequences of 100 new cDNA clones from brain
X which code for large proteins in vitro.";
L DNA Res. 7:143-150(2000).
R EMBL; AB040912; BAA96003.2;
R InterPro; IPR003659; Plexin-like.

```

```

DR Pfam; PF01403; Sema;
DR SMART; SM00423; PSI;
DR SMART; SM00630; Sema; 1.
WK Hypothetical protein.
PT NON TER.
SQ SEQUENCE 1022 AA; 114372 MW; BE4FBD5EA02C69C4 CRC64;

Query Match 42.5%; Score 2314; DB 4; Length 1022;
Best Loca. Similarity 45.4%; Pred. No. 3,5e-179;
Matches 494; Conservative 155; Mismatches 292; Indels 148; Gaps 26;

QY 1 MRSZALLYFTLL---HPAGAGPEDSEPISSISGNTYKQY:PVFVGHKPGRNTCRRHLD 57
DE 12 MRFVLLCAYLILLMVSLRAVSPFEDDEPLNTVDYHSRQYVFRG-RPSGNEQC-HRED 69
CY 59 TCMIMXNGTLYIAARDHIVVDJTSHTBIVCSKLTWKSKRQADVETCRMKKIKDEC 117
DE 70 FQMLCKIRDTLYIAGROQVTVNLNEMPKTEV:PNKALT:WRSQQRENCAYKGRKHKDEC 129
QY 118 HNFIRVLLKNDDALFYCGTNAFYPSCRNYKMTLEPPGDEFGSMARCPYDAXHANVALF 177
DE 130 HNF-KVFPFRNDENVFVCGTNAFNFMCRVYVRLSTLYDGEISGLARCPDARQTNVALF 189
QY 178 AGKLYSATUTDFALDAVYNS:GESPTLRTWKHDSKWLKPEYFVCAVQYGYVEFER 237
DE 190 AGKLYSATVADFASDAVYRSMGCGSALT:KYDSKW:KEPHFLHAIEYGNVYVFER 249
QY 238 EIAVEYNMGKWFERRVAQVCKNDYGGQVLEKQWTSFLKARLNCVSPDSSHFYFNILQ 297
DE 250 EIAVEHNKLGKAVYSVAR:CKNDYGGQVLEKHTWTSFLKARLNCVSPDSSHFYFNILQ 309
QY 278 AVTVIRINGRDVYATFSTPYNS:PGSAVCAVMDL:ASVPTGRFKEQKSPSTWTVPV 357
DE 310 SITDIQINGIPTVGVFTFQNS:IPGSAVCAVMDL:EKVFKGPFKEQKTPGSAWTVP 369
QY 358 DERVKKPRPGCCAGSSSELEVATSNBPDDTLAF:KTHPLNDENAVPS:FNRPWELR:TMVR 417
DE 370 BQVKKPRPGCCAGKIG:AEAYK:SIDPDELTFIKSHPLNCSAVPP:AREPWTK:TRVR 429
QY 418 YELTKIAVDTAAGPYQKHTVYFQSEKXGIIKFLARIGNSGF-LNDSLFLEMSVYNSEK 476
DE 430 YELTAISVDHAGAGYQNTVIFVQSEAGVYLVAK--TSPPSLNDSVLEETEAHNAK 487
QY 477 CSYDGVEDKRIKMGVCLDRASSLYVAFSTVIKVPLOKCRHKKCTIARSDPYCGWI 536
DE 488 CSAENEDKKVLSIQLOKHIALYVAFSSCTIRPLSRCEYVGSCKSCIAERDPYCGWI 547
QY 537 KEGGACSHLSPNSRLT-----FEQDIERGNTDGLGDCHNSFVALNGHSSSL 582
DE 548 SQ-GSCGRVTPGMLLTEDFFAFPHNSAEGYEQDTFGNTAHLCDCRG----- 594
QY 583 LPSTTTSDTAQEGYERGGMLDKHLLDSDPDPJGAVSSHNHQCKGV-RESYLKGH 642
DE 595 -----VRWE--VQSGES----- 604
QY 643 DQGVENVTLALAVILAFMGAVFSGITVYVCD-HRKIVAVYQREKELTHSRGSSVS 701
DE 655 NQWVMVYLITCFAMVVLGAFAGVAVYCYRDMFVKPKI--HKDAESAQSCITSSGS 662
QY 702 VTKLSGLFG---DTQSKDKRKPPEALTFHMHNGKLATEGNTAKYLKADQHH-LDLALPT 757
DE 663 FAKLNGLFDSEVKEVQNIQSPKYSKLLTSRKLPPNGSDTKSMVMHFGQPELAALPT 722
QY 758 PESTPTLQCKRKPGRGREMERNQNLINACTKQVPMGSPVITPDJPLRASPSH:PSVW 817
DE 723 PESTPVLLHCKLQAMKSHSEKAHG--GASRKETPOFFSPSPPHSP--SHGHIPSAIV 778
QY 818 LP-TQCGYQHE?-----VDQP---KQSEVAQMA:EDQAAATLEYVTHREHLS 860
DE 779 LPNATHDYNTSFSNSNAHKAEEKLQNI:DHPLTKGSSKREHRSVDSRNTL--NLLKHLN 836
QY 861 SKSPN-----HGVN:LVENLDSL---PPKVPQREASLGPFGASLSQTGLSKSLEM 906

```

837 0PNSPKA.WGDIQXHCXVLDKXGSMSEVIFKVTNREASVSPSCLPKNSPTXPCV 894  
838 HESSY---GVYKSYFTNUTFSHCATTCKENNTNSSSHLGRNOSFQED-NEPFA 962  
839 PTPGVEMTSLERQGYHKNSQF-HSISAVEK-NLKSPIGVLGRQPSMNEGUMPTT 954  
840 PBYVDSICVHSSQFSQAVTVSRQPSLNAYNSLT-----SSGLKETSLSKPEVPRKSF 1014  
841 GAKVGIQ-----GTPSVHLQPSLSPQSSVTSXGTLPTSLKTESLKEVPRKSF 1007  
842 APJSTSMXP 1025  
843 VPQTFSPRP 1016

Search completed: October 23, 2003, 17:12:38  
Job time : 65 secs



F 26-NOV-1999; 99WO-EP09215.  
X 26-NOV-1998; 98EP-C122441.  
X (PLAC ) MAX PLANCK GES FORDERUNG WISSENSCHAFTEN.  
X Behl C, Kiostermann A;  
X WPI; 2000-400065/34.  
X N-PSDB; AAD0223.  
X Nucleic acid coding for human semaphorin 6A-1 used as diagnostic agent,  
X therapeutic agent, for modulating immune system, in gene therapy or for  
X effecting differentiation, cytoskeletal stabilization and/or plasticity  
X  
X Example 1; Page 29-33; 53pp; English.  
X The present sequence is a transmembranous human semaphorin  
X 6A-1 (HSA:SEMA6A-1) which is involved in neuronal development and  
X regeneration mechanisms during apoptosis. Semaphorin is a family of  
X proteins displaying secreted or transmembrane-based repulsive guidance  
X cues critically involved in neuronal development. The present sequence  
X was isolated from human i-ZAP Express cDNA library which was screened  
X using a PCR fragment amplified from human neuroblastoma cell line  
X SK-N-MC cDNA. The HSA:SEMA6A-1 protein contains a tyrosin-like core  
X that selectively binds to members of Efra/VASP protein family especially  
X Evi. Expression of HSA:SEMA6A-1 is highest in embryonic brain and  
X kidney and moderate in lung. The present sequence is useful as diagnostic  
X and therapeutic agents, for modulating the immune system, in gene  
X therapy, for effecting differentiation, cytoskeletal stabilisation  
X and plasticity.  
X Q Sequence 1030 AA:  
Query Match 100.0%; Score 5450; DB 21; Length 1030;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1030; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Y 1 MRSEALLLYFTLLHFAAGAGFPDESEPISSIGNYTKQYFVGHKFGKNTTQHPIDICW 60  
b 1 MRSEALLLYFTLLHFAAGAGFPDESEPISSIGNYTKQYFVGHKFGKNTTQHPIDICW 60  
Y 61 IYINMGTYAAROHITVDITSHTEELVCSKLTWKSQCAQVCEYKXGKUCSCHNF 120  
b 61 IYINMGTYAAROHITVDITSHTEELVCSKLTWKSQCAQVCEYKXGKUCSCHNF 120  
Y 121 IKVLLKKNDDALFYCGTNAFNSDNYKMTLEFPDGFSGMARCPYDAKHVALEPAC 180  
b 121 IKVLLKKNDDALFYCGTNAFNSDNYKMTLEFPDGFSGMARCPYDAKHVALEPAC 180  
Y 181 KLYSATVTDPLAIDAVIYASLEGPTRTVHSDSKWKKEPYVCAVDYGYVYFFPREIA 240  
b 181 KLYSATVTDPLAIDAVIYASLEGPTRTVHSDSKWKKEPYVCAVDYGYVYFFPREIA 240  
Y 241 VEYNTNGKVPFPRVAQVCKNDMGSCQVLEKQWTSFLKARLNGSVPGDSHFYFNILCAVT 300  
b 241 VEYNTNGKVPFPRVAQVCKNDMGSCQVLEKQWTSFLKARLNGSVPGDSHFYFNILCAVT 300  
Y 301 DVIRINGRDVVLATFTSYNIPISAVCAVDMLDIAVFTGRFKEQKSPDSTWTPVDPER 360  
b 301 DVIRINGRDVVLATFTSYNIPISAVCAVDMLDIAVFTGRFKEQKSPDSTWTPVDPER 360  
Y 361 VPKPBPCCAGSSSLERVATSNRPDDTLNFKTHPMDGAVPSIFNRPFLATMYRYRL 420  
b 361 VPKPBPCCAGSSSLERVATSNRPDDTLNFKTHPMDGAVPSIFNRPFLATMYRYRL 420  
Y 421 TKIAVDTAAGYQNHVTVFLGSEKGIILKFLARLNGSFLNDSFLLEMSVYNSEKCSYD 480  
b 421 TKIAVDTAAGYQNHVTVFLGSEKGIILKFLARLNGSFLNDSFLLEMSVYNSEKCSYD 480  
Y 481 GVEDKRMINGMQLDRASSLYVAFTTCVIKVPVLCGRCHGKCKXTCTIASRDPYCGWIKEGG 540  
b 481 GVEDKRMINGMQLDRASSLYVAFTTCVIKVPVLCGRCHGKCKXTCTIASRDPYCGWIKEGG 540

Db 481 GVEDKRMINGMQLDRASSLYVAFTTCVIKVPVLCGRCHGKCKXTCTIASRDPYCGWIKEGG 540  
Qy 541 ACSHLSFNSRLTTEQDIERNITDGLGDCHNSFVALKNGHSSSLPSTTTSDSTAQEGYSR 600  
Db 541 ACSHLSFNSRLTTEQDIERNITDGLGDCHNSFVALKNGHSSSLPSTTTSDSTAQEGYSR 600  
Qy 601 GGMLDWKHLSDSPDTPGLGAVSSHHNDDKKGVIRSYLKGHDQLVFTLLA:AVTLAFV 660  
Db 601 GGMLDWKHLSDSPDTPGLGAVSSHHNDDKKGVIRSYLKGHDQLVFTLLA:AVTLAFV 660  
Qy 661 MGAVFSGITVYVCVCHRRKQVAVVQKKEKELTHSRGSMSSVTKLSLFGDTOSKDPKPE 720  
Db 661 MGAVFSGITVYVCVCHRRKQVAVVQKKEKELTHSRGSMSSVTKLSLFGDTOSKDPKPE 720  
Qy 721 AILPPLVHNGKLATPGNTAKYLKADQHHLDLALPTPESTPTLQQRKRPSPRSREWRN 780  
Db 721 AILPPLVHNGKLATPGNTAKYLKADQHHLDLALPTPESTPTLQQRKRPSPRSREWRN 780  
Qy 781 CNLINACTKYPKMGSPVPTDLPJLRASPSHIFSVVVLPTQOQYQHEVYVQPKYSEVAQ 840  
Db 781 CNLINACTKYPKMGSPVPTDLPJLRASPSHIFSVVVLPTQOQYQHEVYVQPKYSEVAQ 840  
Qy 841 VAIJEDQATILEYKTIKEHLSKSPNHGVN:VENLDSLPPKVPQREASLGPFGASLSQTGL 900  
Db 841 VAIJEDQATILEYKTIKEHLSKSPNHGVN:VENLDSLPPKVPQREASLGPFGASLSQTGL 900  
Qy 901 SKRLEMHSSSYGVYVYASPTNSL:TRSHQATTLKKNNTNSGSHLSRNQSGRGDNP 960  
Db 901 SKRLEMHSSSYGVYVYASPTNSL:TRSHQATTLKKNNTNSGSHLSRNQSGRGDNP 960  
Qy 961 PAPQKVDSDIWHSSQPSQAVTVSRQPS:NAVNSL:TRSGLKRTPSLKPDVPEKPSFAPLS 1020  
Db 961 PAPQKVDSDIWHSSQPSQAVTVSRQPS:NAVNSL:TRSGLKRTPSLKPDVPEKPSFAPLS 1020  
Qy 1021 TSMKPNDACT 1030  
Db 1021 TSMKPNDACT 1030  
RESULT 2  
ID AAB95731 standard; Protein: 975 AA.  
XX AAB95731;  
XX 07-JUN-2002 (first entry)  
XX Human CJ45-1 protein sequence SEQ ID 161.  
XX Human; secreted protein; nutrition; cytokine modulator; proliferation;  
XX differentiation; immune system modulator; tissue growth; chemotactic;  
XX haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;  
XX haematopoiesis.  
XX Homo sapiens.  
XX WC20011998A-1.  
XX 22-MAR-2001.  
XX 14-SEP-2003; 2000MC-US251135.  
XX 17-SEP-1999; 99JS-0398829.  
XX (GENY ) GENETICS INST. INC.  
XX Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;  
XX Merberg D, Treacy M, Bowman MR, Spaulding V, Agostino MJ;  
XX WPI; 2001-244801/25.  
XX N-PSDB; AAF98469.  
XX Isolated nucleic acids encoding polypeptides, useful for modulating

e.g. cytokine and cell proliferation/differentiation activity, the immune system and hematopoiesis regulating activity.

Disclosure: Page 487-490; 557pp; English.

Human cDNA clones represented in AAF98374 - AAF98489 encode secreted proteins AAF96667 - AAF98075. The cDNA clones are isolated from various tissue types, and may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate protein expression. The polypeptides and nucleic acids may be used as nutrients or to modulate cytokine and cell proliferation/differentiation activity and may also be involved in regulation of the immune system. The cDNA sequences, proteins, their agonists and/or antagonists exhibit hematopoiesis regulating activity, tissue growth activity, activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombotic activity; receptor/ligand activity; anti-inflammatory activity; hematopoiesis activity; cachectin/tumor suppressor activity; and/or tumour inhibition activity. Included in the invention are probes represented in AAF98490 - AAF98572 which are specific for the cDNA clones encoding the secreted proteins.

Sequence 975 AA:

```

Query Match      93.6%; Score 5103.5; DB 22; Length 975;
Best Local Similarity 94.4%; Pred. No. 2;
Matches 972; Conservative 0; Mismatches 3; Indels 55; Gaps 1;

1  MSSEALLLYFTLLHFAGAGPEDESEPTISGHNYTKQYVFVGHKPGKNTQTHRLDTCM 60
|||||
1  MSSEALLLYFTLLHFAGAGPEDESEPTISGHNYTKQYVFVGHKPGKNTQTHRLDTCM 60

61  IMINNGTLYIAARDHIYTVDTISHTTEIYCSKLTWKSROAEVDTQYKSKHDCENHP 120
|||||
61  IMINNGTLYIAARDHIYTVDTISHTTEIYCSKLTWKSROAEVDTQYKSKHDCENHP 120

121  IKVLLKKNDJALFVCGTNAFNPCSNYKMDTEPEGDEFSGMARCPYDAKANYALPADG 180
|||||
121  IKVLLKKNDJALFVCGTNAFNPCSNYKMDTEPEGDEFSGMARCPYDAKANYALPADG 180

181  KLYSATVTDFAIDAIVIRSLGESPTRTVTHCSKWLKEPYVQAVDYGDIYFFFEIRA 240
|||||
181  KLYSATVTDFAIDAIVIRSLGESPTRTVTHCSKWLKEPYVQAVDYGDIYFFFEIRA 240

241  VEYNTYGVKVPFPRVACVKNYCGGSGORVLEKQWTSFLKAPLNCSPGGSHFVNTLOAVT 300
|||||
241  VEYNTYGVKVPFPRVACVKNYCGGSGORVLEKQWTSFLKAPLNCSPGGSHFVNTLOAVT 300

301  DVIRINGROWLATFTSTPYNLS:PGSAVCAVYDMLD:ASVFTGSEKQKSPDSTWTPVPLEP 360
|||||
301  DVIRINGROWLATFTSTPYNLS:PGSAVCAVYDMLD:ASVFTGSEKQKSPDSTWTPVPLEP 360

361  VPKPRGCCAGSSSLERYATSNFEPDQTNFIKTHPLNDEAVPS:FNENFELRTWYVRL 420
|||||
361  VPKPRGCCAGSSSLERYATSNFEPDQTNFIKTHPLNDEAVPS:FNENFELRTWYVRL 420

421  TKIADVTAAGVQNHVTVFVJGSEKGIKFKFARIQNSGFNDLSLPLEKSVYVSEKQSYD 480
|||||
421  TKIADVTAAGVQNHVTVFVJGSEKGIKFKFARIQNSGFNDLSLPLEKSVYVSEKQSYD 480

481  GVECKIKNGQDRASSSLVYAFSTCVIKVPLGRERHKGKCKTKIARDPYCYGMVKEGG 540
|||||
481  GVECKIKNGQDRASSSLVYAFSTCVIKVPLGRERHKGKCKTKIARDPYCYGMVKEGG 540

541  ACSHLSPNRLTFEOD:ERGVTDGLGOCNHSFVALNCHSSSLDPS:TTTSCDSTAGQGVSR 600
|||||
541  ACSHLSPNRLTFEOD:ERGVTDGLGOCNHSFVALNCHSSSLDPS:TTTSCDSTAGQGVSR 600

601  GGLDWKHLLEDSPDSTPLGAVSSHQKKGKGVIRESYLKGHDQLVPYVTLIAIVILAFV 660
|||||
601  GGLDWKHLLEDSPDSTPLGAVSSHQKKGKGVIRESYLKGHDQLVPYVTLIAIVILAFV 660

661  NGAVFSGITVYVCDHRRKDVAVVQRKEKELTHSRRGSMSSVTKLSGLFGDTQSKDKPKE 720
|||||
661  NGAVFSGITVYVCDHRRKDVAVVQRKEKELTHSRRGSMSSVTKLSGLFGDTQSKDKPKE 720

```

```

Db      606  MGAVFSGITVYVCDHRRKDVAVVQRKEKELTHSRRGSMSSVTKLSGLFGDTQSKDKPKE 665
Qy      721  AILTPXNHGKATPONTAKMLKADQHHLDLTALPTPESTPTLQOKRXPDSRGSREWERN 780
Db      666  AILTPXNHGKATPONTAKMLKADQHHLDLTALPTPESTPTLQOKRXPDSRGSREWERN 725
Qy      781  QNLINACTKMPSPKSPV:PTDLPLRASPSH:PSVVLVLPITQGYCHEYVDCPKSEVAC 840
Db      726  QNLINACTKMPSPKSPV:PTDLPLRASPSH:PSVVLVLPITQGYCHEYVDCPKSEVAC 785
Qy      841  MALEDOAAATLEVNTIKERHSSKSPNKGVLNENLDSLPKVPQREASLGPPOCASLQTL 900
Db      786  MALEDOAAATLEVNTIKERHSSKSPNKGVLNENLDSLPKVPQREASLGPPOCASLQTL 845
Qy      901  SXPLEMHSSGVGVYKRSYPTNSLTRSHOATLKNNTNSNSHLSNQSFGGDNTP 960
Db      846  SXPLEMHSSGVGVYKRSYPTNSLTRSHOATLKNNTNSNSHLSNQSFGGDNTP 905
Qy      961  PAPCRVDSIQVHSSQPSGCAVTVSRQPS:NAVNSLTRSGLKXETPSLKPDVPPKPSFAPLS 1020
Db      906  PAPCRVDSIQVHSSQPSGCAVTVSRQPS:NAVNSLTRSGLKXETPSLKPDVPPKPSFAPLS 965
Qy      1021  TSMKPNDACT 1030
Db      966  TSMKPNDACT 975

RESULT ?
AAW64221
ID  AAW64221 standard; Protein: 974 AA.
XX
AC  AAW64221;
XX
DT  06-OCT-1998 (first entry;
XX
DE  Human secreted protein from clone C0145_1.
XX
KW  Secreted protein; human fetal brain; nutrition; cytokine; stimulant;
KW  cell proliferation; differentiation; immune system; suppressor; ligand;
KW  regulator; hematopoiesis; tissue growth; activin; inhibin; haemostatic;
KW  chemotaxis; chemokinetic; thrombosis; receptor; cadherin; tumour;
KW  anti-inflammatory.
XX
OS  Homo sapiens.
XX
XX  W09827205-A2.
XX
XX  25-JUN-1998.
XX
XX  17-DEC-1997; 97MO-US23330.
XX
XX  16-DEC-1997; 97US-0991872.
XX  18-DEC-1996; 96US-0769192.
XX  13-JAN-1997; 97US-0783431.
XX
XX  GENEV : GENETICS INST INC.
XX
XX  Agostino MJ, Jacobs K, Lavallie ER, McCoy JX, Merberg D;
XX  Racine SA, Spaulding V, Treacy N;
XX
XX  WP: 1996-362774/31.
XX  N-PSDS; AAV44295.
XX
XX  New polynucleotides and secreted proteins - obtained from human
XX  foetal brain, human adult testes, human adult brain and human adult
XX  salivary gland cDNA libraries
XX
XX  Claim 173; Page 71-74; l10pp; English.
XX
XX  This sequence represents a novel secreted protein from clone C0145_1
XX  isolated from a human fetal brain cDNA library. This protein has
XX  applications for nutritional use, cytokine and cell
XX  proliferation/differentiation activity, immune stimulating or

```

C suppressing activity, hematopoiesis regulating activity, tissue growth  
C activity, activin/inhibin activity, chemotactic/chemokinetic activity,  
C haemostatic and thrombotic activity, receptor/ligand activity,  
C anti-inflammatory activity, cadherin/tumour invasion suppressor activity,  
C tumour inhibition activity and other activities.

X Query Match 93.4%; Score 5089; DB 19; Length 974;  
X Best Local Similarity 94.3%; Evid. No. 3;  
X Matches 971; Conservative 0; Mismatches 3; Indels 56; Gaps 2;

1 MRSEALLYFTLLHFAAGPESEPEISISNGNYTKQYVVFVGHKPGRNTHQRLDQM 60  
b 1 MRSEALLYFTLLHFAAGPESEPEISISNGNYTKQYVVFVGHKPGRNTHQRLDQM 60  
/ 61 IMXNGTLYAARDHIYTVVIDSHTEEEIVCSKLTWKSRQADVTCRMKGKHKECHNF 120  
b 61 IMXNGTLYAARDHIYTVVIDSHTEEEIVCSKLTWKSRQADVTCRMKGKHKECHNF 119  
/ 121 IKVLLKXNDALFYCGTNAPNPSGRNKKOTLDFPQDEFSGMARQPYDAHANVALFADQ 180  
b 121 IKVLLKXNDALFYCGTNAPNPSGRNKKOTLDFPQDEFSGMARQPYDAHANVALFADQ 179  
/ 181 KLYSATVTDPLAIDAVYRSLGSEPTRTYKHCSKCLKERYPCANLYGQYVFFPREIA 240  
b 181 KLYSATVTDPLAIDAVYRSLGSEPTRTYKHCSKCLKERYPCANLYGQYVFFPREIA 239  
/ 241 VEYTMGKVFVFAQCKNDKMSGRVJSEKQWTSFLKAPLNGSVFGDGHFFNLIQAVT 300  
b 241 VEYTMGKVFVFAQCKNDKMSGRVJSEKQWTSFLKAPLNGSVFGDGHFFNLIQAVT 299  
/ 301 DVIRINGEDVLAFTSPYNSIPGSAVCATQMDLQASVFTQRPKEKSGCDSTWTVPEER 360  
b 301 DVIRINGEDVLAFTSPYNSIPGSAVCATQMDLQASVFTQRPKEKSGCDSTWTVPEER 359  
/ 361 VPKPRPGCCAGSSSLEAYATSNPEFDTLNFIKTHPLMDKAVPSIFNRKMLRTWVEYRI 420  
b 361 VPKPRPGCCAGSSSLEAYATSNPEFDTLNFIKTHPLMDKAVPSIFNRKMLRTWVEYRI 419  
/ 421 TKIADVTAGPYQWHTVVFLOSGNGIILKELARIGNSGFLNDSIFLEKSVYNSEKQYD 480  
b 421 TKIADVTAGPYQWHTVVFLOSGNGIILKELARIGNSGFLNDSIFLEKSVYNSEKQYD 479  
/ 481 GVEKRLVWGLODASSSLVAFSTCVKPLGRGBHOKKXKTQASRQPCGNTKEG 540  
b 481 GVEKRLVWGLODASSSLVAFSTCVKPLGRGBHOKKXKTQASRQPCGNTKEG 539  
/ 541 ACSHLSFNSRLTFFQDIERGNTDCLGSCNHSFVALKACHSSELPSTTTSDSKAQEYER 600  
b 541 ACSHLSFNSRLTFFQDIERGNTDCLGSCNHSFVALKACHSSELPSTTTSDSKAQEYER 599  
/ 545 ACSHLSFNSRLTFFQDIERGNTDCLGSCNHSFVALKACHSSELPSTTTSDSKAQEYER 575  
b 545 ACSHLSFNSRLTFFQDIERGNTDCLGSCNHSFVALKACHSSELPSTTTSDSKAQEYER 574  
/ 546 GGMLDKXILLDSPESTPLGAVSHKHQKGVHESYLKGHPLQVPLLAIAVLAIAV 660  
b 546 GGMLDKXILLDSPESTPLGAVSHKHQKGVHESYLKGHPLQVPLLAIAVLAIAV 659  
/ 576 -----GVRESYLKGHPLQVPLLAIAVLAIAV 664  
b 576 -----GVRESYLKGHPLQVPLLAIAVLAIAV 663  
/ 661 MGAVFSGITVYVCVCHRRKDVAVYORKEKSLTHSRGSMSSVTKLGLFQDTCQKDPKE 720  
b 661 MGAVFSGITVYVCVCHRRKDVAVYORKEKSLTHSRGSMSSVTKLGLFQDTCQKDPKE 719  
/ 665 MGAVFSGITVYVCVCHRRKDVAVYORKEKSLTHSRGSMSSVTKLGLFQDTCQKDPKE 780  
b 665 MGAVFSGITVYVCVCHRRKDVAVYORKEKSLTHSRGSMSSVTKLGLFQDTCQKDPKE 779  
/ 721 AILPLMHNGKLTATPGNTAKMLKACHHLDLIALTPPESTPTLQQRKXPSGSRWERN 780  
b 721 AILPLMHNGKLTATPGNTAKMLKACHHLDLIALTPPESTPTLQQRKXPSGSRWERN 779  
/ 665 AILPLMHNGKLTATPGNTAKMLKACHHLDLIALTPPESTPTLQQRKXPSGSRWERN 724  
b 665 AILPLMHNGKLTATPGNTAKMLKACHHLDLIALTPPESTPTLQQRKXPSGSRWERN 723  
/ 781 QNLINACTKOMP2YXGSPVPTDGLFASPSHISVWVLPITQOQYCHEYVDQPKSEVAC 840  
b 781 QNLINACTKOMP2YXGSPVPTDGLFASPSHISVWVLPITQOQYCHEYVDQPKSEVAC 839  
/ 725 QNLINACTKOMP2YXGSPVPTDGLFASPSHISVWVLPITQOQYCHEYVDQPKSEVAC 784  
b 725 QNLINACTKOMP2YXGSPVPTDGLFASPSHISVWVLPITQOQYCHEYVDQPKSEVAC 783  
/ 841 MAJEDQAAHTEYKTIKEHLSKSNHGNLVENLDSLP7PKVPQREASLAPPASLSQGL 900  
b 841 MAJEDQAAHTEYKTIKEHLSKSNHGNLVENLDSLP7PKVPQREASLAPPASLSQGL 899  
/ 785 VALEQCAATLETKIKENFSKSNHGNLVENLDSLP7PKVPQREASLAPPASLSQGL 844  
b 785 VALEQCAATLETKIKENFSKSNHGNLVENLDSLP7PKVPQREASLAPPASLSQGL 843  
/ 901 SKRLEMHHSYGYVYKRSYPTNSLTRSHQATTLKNNNTNSNSHLSRNSQSGRGDNP 960  
b 901 SKRLEMHHSYGYVYKRSYPTNSLTRSHQATTLKNNNTNSNSHLSRNSQSGRGDNP 959

Db 845 SKRLEMHHSYGYVYKRSYPTNSLTRSHQATTLKNNNTNSNSHLSRNSQSGRGDNP 904  
Qy 961 PAFQVDSIQVHSSQSGGAVTVSRQPSLNAYNSLTRSGLRKTFPSLKQDVPKPSFAPLS 1020  
Dt 905 PAFQVDSIQVHSSQSGGAVTVSRQPSLNAYNSLTRSGLRKTFPSLKQDVPKPSFAPLS 964  
Qy 1021 TSMKPNDACT 1030  
Db 965 TSMKPNDACT 974

RESULT 4  
AAB2303C  
ID AAB2303C standard; Protein: 939 AA.  
XX  
AC AAB2303C;  
XX  
XX  
DT 16-JAN-2001 (first entry)  
XX  
XX Human semaphorin: protein-like splice variant, SECX 2864933-1.  
XX  
XX SECX protein; human; secreted; membrane-associated; cancer;  
XX proliferation regulator; differentiation regulator; non-malignant tumour;  
XX immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;  
XX infection; inflammatory disorder; arthritis; haematopoietic disorder;  
XX skin disorder; cardiovascular disorder; atherosclerosis; restenosis;  
XX neurological disease; Alzheimer's disease; trauma; wound;  
XX spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;  
XX anti-HIV; anti-inflammatory; anti-arthritis; antiarteriosclerotic;  
XX neuroprotective; vulnery; anti-allergic; antimicrobial; cardiac;  
XX dermatological; gene therapy.  
XX  
XX Homo sapiens.  
XX  
XX WC0200053742-A2.  
XX  
XX 14-SEP-2000.  
XX  
XX 09-MAR-2000; 2000W0-JSC6280.  
XX  
XX 09-MAR-1999; 99US-0123667.  
XX  
XX 06-MAR-2000; 2000US-0123667.  
XX  
XX (CUBA-1) CURAGEN CORP.  
XX  
XX Shimkats RA;  
XX  
XX MPI; 2000-5943-B/56.  
XX  
XX N-PSDB; AAA93617.  
XX  
XX Novel human membrane associated or secreted polypeptides and  
XX polynucleotides useful for diagnosis, prevention and treatment of  
XX pathological states such as cancer, immune, cardiovascular and  
XX neurological disorders  
XX  
XX Claim 1; Fig 2; 151pp; English.  
XX  
XX Sequences AAB23029-B23049 represent human SECX proteins. The SECX  
XX proteins of the invention are either secreted or membrane-associated  
XX proteins and act as regulator of cellular proliferation and  
XX differentiation. SECX proteins or nucleotides are useful for diagnosing  
XX the presence of, or predisposition to, a disease associated with altered  
XX levels of SECX proteins and nucleotides. The SECX proteins are also  
XX useful to screen compounds that modulate SECX activity or expression. The  
XX interaction of a SECX protein with other cellular proteins may be useful  
XX to modulate the activity of a partner protein, cellular proliferation,  
XX cellular differentiation and cell survival. SECX nucleotides are useful  
XX for the recombinant expression of SECX protein. They may also be used to  
XX SECX mRNA or genetic lesions in the SECX gene. They may also be used to  
XX modulate SECX expression (e.g., using antisense oligonucleotides). SECX  
XX nucleic acid sequences are also useful for identifying a cell or tissue  
XX type in a biological sample, and in forensic biology. SECX primers or

QY	841	MALEQAAATLSYKTIKEHLKSKSNHGVN:AVENLDSJFFKVPQREASLGPFGASLSCTGL	900
DE	841	MALEQAAATLSYKTIKEHLKSKSNHGVN:AVENLDSLPPKVPQREASLGPFGASLSCTGL	900
QY	901	SKRLNHHSSSGYGVYKPSYPTNSLTRSHQAT	932
DE	901	SKRLNHHSSSGYGVYKPSYPTNSLTRSHLIT	932

RESULT 5	
AAB2303:	
ID	AAB2303; standard; Protein; 694 AA.
XX	
AC	AAB2303;
XX	
C-	16 JAN-2004 (first entry)
XX	
DE	Human semaphorin protein-like splice variant, SECC 2864933-2.
XX	
SW	SECC protein; human; secreted; membrane-associated; cancer;
KW	proliferation regulator; differentiation regulator; non-malignant tumor;
KW	immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;
KW	infection; inflammatory disorder; arthritis; haematopoietic disorder;
KW	skin disorder; cardiovascular disorder; arteriosclerosis; restenosis;
KW	neurological disease; Alzheimer's disease; trauma; wounding;
KW	spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;
KW	anti-HV; anti-inflammatory; antitubercic; antiarteriosclerotic;
KW	neurprotective; vulnerrary; antiallergic; antimicrobial; cardiant;
KW	dermatological; gene therapy.

XX	CS	Homo sapiens.
XX	XX	
XX	PN	WQ200053742-A2.
XX	XX	
XX	PD	14-SEP-2000.
XX	XX	
XX	PF	03-MAR-2000; 2000MO-US06220.
XX	XX	
XX	PR	05-MAR-1993; 99US-0123667.
XX	PR	08-MAR-2000; 2000US-0123667.
XX	XX	
XX	FA	'CURA-1' CURASEN CORP.
XX	XX	
XX	FI	Shimkets RA;
XX	XX	
XX	FI	WP1; 2000-594312/56.
XX	CR	N-PSDB; AAA93619.
XX	XX	
XX	XX	Note: human membrane associated or secreted polypeptides and
XX	XX	polynucleotides useful for diagnosis, prevention and treatment of
XX	XX	pathological states such as cancer, immune, cardiovascular and
XX	XX	neurological disorders -
XX	XX	
XX	XX	Claim 1; Fig 3; 15:pp; English.

Sequences AB230203-B23048 represent human SECX proteins. The SECX proteins of the invention are either secreted or membrane-associated proteins and act as regulator of cellular proliferation and differentiation. SECX proteins or nucleotides are useful for diagnosing the presence of, or predisposition to, a disease associated with altered levels of SECX proteins and nucleotides. The SECX proteins are also useful to screen compounds that modulate SECX activity or expression. The interaction of a SECX protein with other cellular proteins may be useful to modulate the activity of a partner protein, cellular proliferation, cellular differentiation and cell survival. SECX nucleotides are useful for the recombinant expression of SECX protein, and may also be used to detect SECX RNA or genomic lesions in the SECX gene. They may also be used to modulate SECX expression (e.g., using antisense oligonucleotides). SECX nucleic acid sequences are also useful for identifying a cell or tissue type in a biological sample, and in forensic biology. SECX primers or probes are useful for detecting the presence of SECX nucleotides and for screening tissue cultures for contamination. Diseases that may be treated or prevented using SECX proteins or nucleotides include cancer (e.g.,



C Colorectal carcinoma, prostate cancer, benign tumors, immune disorders  
C (including autoimmune diseases, transplant rejection, allergies, AIDS),  
C infections, inflammatory disorders, arthritis, hematopoietic disorders,  
C skin disorders, cardiovascular disorders, atherosclerosis, restenosis,  
C neurological diseases (e.g., Alzheimer's disease), trauma (e.g.,  
C surgical or traumatic wounds, spinal cord injury), and skeletal  
C disorders.  
X  
C Sequence 894 AA:  
Query Match 84.2%; Score 4590.5; DB 2; Length 894;  
Best Local Similarity 93.7%; Pred. No. 0;  
Matches 873; Conservative 2; Mismatches 2; Indels 55; Gaps 1;  
Y 1 MESSALLLYPTLAFAGAGPEDEPISISHGNYTKQYVFQKFKGRNTTQSHRLDLCM 60  
E 1 YRSALJJYFTLLAFAGAGPEDEPIS:SHGNYTKQYVFQKFKGRNTTQSHRLDLCM 60  
Y 61 IMINGTJYNAARDHITVDITSHTEBEEYCSKLLTWKSPQADVTCTCKKGFHDECHNF 120  
D IMINGTJYNAARDHITVDITSHTEBEEYCSKLLTWKSPQADVTCTCKKGFHDECHNF 120  
Y 121 IKVILKKKDALFVCGNAPNESLNNKVTLEFFGDEPSGMARCPYDAKHANVALPADG 180  
D IKVILKKKDALFVCGNAPNESLNNKVTLEFFGDEPSGMARCPYDAKHANVALPADG 180  
Y 181 KYLSATVDFLDAIDAVYRSLGESPTLTVKHDSKMLKEPYFVQAYDGYDYFFPREIA 240  
b KYLSATVDFLDAIDAVYRSLGESPTLTVKHDSKMLKEPYFVQAYDGYDYFFPREIA 240  
Y 241 VEYTMKGVNFPVQACVCKNDGSGORVLEKQNTSELKARLNCVPGDSHFVENILQANT 300  
D VEYTMKGVNFPVQACVCKNDGSGORVLEKQNTSELKARLNCVPGDSHFVENILQANT 300  
Y 301 DVIRINDVYLATFTFYNS:PGSAVCAYDMLASVFTGTFKEKGKQSDSTWTPVEDER 360  
D DVIRINDVYLATFTFYNS:PGSAVCAYDMLASVFTGTFKEKGKQSDSTWTPVEDER 360  
Y 361 VPKRPGCCAGSSLEAYATSNFPPDITLFIKTHPLNDKAVESIENRPFJBTMHPYSL 420  
D VPKRPGCCAGSSLEAYATSNFPPDITLFIKTHPLNDKAVESIENRPFJBTMHPYSL 420  
Y 421 TKIADVTANGPYQNTVVFQSEKGIILKELARLGNISCHLNDLFLFVMSVNSKXCSVD 480  
D TKIADVTANGPYQNTVVFQSEKGIILKELARLGNISCHLNDLFLFVMSVNSKXCSVD 480  
Y 481 QVEDKRVGMQDRASSJYVAFSTCVIKVPLGCEPHGKCKTCTIAROPYGCMVKEGG 540  
D QVEDKRVGMQDRASSJYVAFSTCVIKVPLGCEPHGKCKTCTIAROPYGCMVKEGG 540  
Y 541 ACSNLSNSRLTFQRIERGTGCGGCHNSFALNCHASSLIPSTTSDSTACQCYVER 600  
D ACSNLSNSRLTFQRIERGTGCGGCHNSFALNCHASSLIPSTTSDSTACQCYVER 600  
Y 601 GGMLOWKILLSPDSTPLGAVSSHHNQKNGVIRESYLKGHPQCVFVTLAIAVLAAY 660  
D GVMRFSYLAGHQQVPTVTLAIAVLAAY 665  
Y 661 MGAVFSGITVYVCDHRKDVAVYCRKEKELTSPRSGMSVTKLSGLFCTOSKPKPE 720  
b MGAVFSGITVYVCDHRKDVAVYCRKEKELTSPRSGMSVTKLSGLFCTOSKPKPE 725  
Y 721 AILTPDMHNGKLATPGNTAKMLIADQHHLDLTALPTPESTFTLQQRKPSHSGSREWRN 780  
b AILTPDMHNGKLATPGNTAKMLIADQHHLDLTALPTPESTFTLQQRKPSHSGSREWRN 785  
Y 781 QNLINACTKMPKPGSPVPTDLFLRASPSPHPSVVLPTTQQCYCHYYVDQPKYSEVAC 840  
b QNLINACTKMPKPGSPVPTDLFLRASPSPHPSVVLPTTQQCYCHYYVDQPKYSEVAC 845  
Y 841 MALEDCAATLEYKTIKEHLSSKSNHGNLVEN:LSLPPKVPQBEASLGPQASLSQTGL 900  
D MALEDCAATLEYKTIKEHLSSKSNHGNLVEN:LSLPPKVPQBEASLGPQASLSQTGL 945

QY 901 SKLEMHGSSVGVYKRSVPNTSLTRSHOAT 932  
DB 846 SKLEMHGSSVGVYKRSVPNTSLTRSHOAT 877  
RESULT 6  
AAB95139  
ID AAB95139 standard; Protein: 699 AA.  
XX AAB95139;  
AC AAB95139;  
XX 26-JUN-2001 (first entry;  
XX Human protein sequence SEQ ID NO:17154.  
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
XX Homo sapiens.  
XX EP:074617-A2.  
XX 07-FEB-2001.  
XX 28-JUN-2000; 2000EP-0116:26.  
XX 29-JUL-1999; 99JP-0248026.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0193767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX (HELIX) HELIX RES INST.  
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
WPI: 2001-318749/34.  
XX Primer sets for synthesizing polynucleotides, particularly the 5602  
full-length cDNAs defined in the specification, and for the detection  
and/or diagnosis of the abnormality of the proteins encoded by the  
full-length cDNAs -  
PS Claim 3; SEQ ID 17154; 2337pp + CD ROM; English.  
XX The present invention describes primer sets for synthesizing 5602  
full-length cDNAs defined in the specification. Where a primer set  
comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
to the complementary strand of a polynucleotide which comprises one of  
the 5602 nucleotide sequences defined in the specification, where the  
oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
of an oligonucleotide comprising a sequence complementary to the  
complementary strand of a polynucleotide which comprises a 5'-end  
sequence and an oligonucleotide comprising a sequence complementary to a  
polynucleotide which comprises a 3'-end sequence, where the  
oligonucleotide comprises at least 15 nucleotides and the combination of  
the 5'-end sequence/3'-end sequence is selected from those defined in  
the specification. The primer sets can be used in antisense therapy and  
in gene therapy. The primers are useful for synthesizing polynucleotides,  
particularly full-length cDNAs. The primers are also useful for the  
detection and/or diagnosis of the abnormality of the proteins encoded by  
the full-length cDNAs. The primers allow obtaining of the full-length  
cDNAs easily without any specialised methods. AAB03166 to AAB13628 and  
AAB13631 to AAB18742 represent human cDNA sequences; AAB92446 to  
AAB95893 represent human amino acid sequences; and AAB13629 to AAB13632  
represent oligonucleotides, all of which are used in the exemplification  
of the present invention.  
XX Sequence 699 AA;  
Query Match 97.5%; Score 3677; DB 2; Length 699;  
Best Local Similarity 100.0%; Pred. No. 1.6e-310;

Matches 699; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

332 MLDIASVTGKEKEKSPDSTWTPDQVPVPRPGCCACSSSLERYATSNRPEDTLNF 391  
|||||  
1 MLDIASVTGKEKEKSPDSTWTPDQVPVPRPGCCACSSSLERYATSNRPEDTLNF 60  
|||||  
392 IKTHPLMDAVSIFNRPFLATVYVYELTKIAVPTAAGPYQWHTWFGSEKGIILKFL 451  
|||||  
61 IKTHPLMDAVSIFNRPFLATVYVYELTKIAVPTAAGPYQWHTWFGSEKGIILKFL 120  
|||||  
452 ARISNGPLNDSLFLEKSVYNSKSCSDGVQDKRMQMLDRASSSLVAFSTGVKVP 511  
|||||  
121 ARISNGPLNDSLFLEKSVYNSKSCSDGVQDKRMQMLDRASSSLVAFSTGVKVP 120  
|||||  
512 LGRCERHKGKCTCTASRPYQWKEGAGSHSUNSLTFEGQIEGQNTDGLGCHNS 571  
|||||  
181 LGRCERHKGKCTCTASRPYQWKEGAGSHSUNSLTFEGQIEGQNTDGLGCHNS 240  
|||||  
572 FVALKGHSSLLFTTTSSTAQEGYERGGMLDKWILLDPESLTFGLAYSSSNHQDKX 631  
|||||  
241 FVALKGHSSLLFTTTSSTAQEGYERGGMLDKWILLDPESLTFGLAYSSSNHQDKX 300  
|||||  
632 GVIRESYLKHDPVPTLALIAVLAPVGAFFSITVYCVDRHRKQDAVYQKKEKL 691  
|||||  
301 GVIRESYLKHDPVPTLALIAVLAPVGAFFSITVYCVDRHRKQDAVYQKKEKL 360  
|||||  
692 THSRRGSSSVTKLSLGGTOSKPKPEALITPLMHNKGLATPNTAKMLIKADQHLE 751  
|||||  
361 THSRRGSSSVTKLSLGGTOSKPKPEALITPLMHNKGLATPNTAKMLIKADQHLE 420  
|||||  
752 LTAJLTPTSTPLOGKRPSGSRERNQNLINACTKMPERMGSVPITPLIRASPSH 811  
|||||  
421 LTAJLTPTSTPLOGKRPSGSRERNQNLINACTKMPERMGSVPITPLIRASPSH 480  
|||||  
812 IPSVVLPITQGGYCHEYVQDPKSEVAQMALEDQAATLEYNTKEHLSSKSPHGNVLY 871  
|||||  
481 IPSVVLPITQGGYCHEYVQDPKSEVAQMALEDQAATLEYNTKEHLSSKSPHGNVLY 540  
|||||  
872 ENLDSPPKVPQREASLGPEGASLSQTLGSKLEMHSSYGVYKSPYNTSLTRSHOA 931  
|||||  
541 ENLDSPPKVPQREASLGPEGASLSQTLGSKLEMHSSYGVYKSPYNTSLTRSHOA 600  
|||||  
932 TTKRNNNTSSKSSLSRNCQSGRGNPPAPQVDSIQVHSSQSGQAVVSGQFSLNA 991  
|||||  
601 TTKRNNNTSSKSSLSRNCQSGRGNPPAPQVDSIQVHSSQSGQAVVSGQFSLNA 660  
|||||  
992 YNSLTRSGIKPTPSUKPDPVPPKSPFAPJSTSMKPRDACT 1030  
|||||  
661 YNSLTRSGIKPTPSUKPDPVPPKSPFAPJSTSMKPRDACT 639  
|||||

RESULT 7

AB23043  
O AAB23043 standard; Protein; 630 AA.  
X C AAB23043;  
X C AAB23043;  
T 16-JAN-2001 (first entry;  
T Human semaphorin protein-like splice variant, SECX pC2.1-2864933.  
E SECX protein; human; secreted; membrane-associated; cancer;  
E proliferation regulator; differentiation regulator; not malignant tumour;  
E immune disorder; autoimmune disease; transplant rejection; allergy; AIDS;  
E infection; inflammatory disorder; arthritis; haematopoietic disorder;  
E skin disorder; cardiovascular disorder; atherosclerosis; restenosis;  
E neurological disease; Alzheimer's disease; atherosclerosis; restenosis;  
E spinal cord injury; skeletal disorder; cytostatic; immunosuppressive;  
E anti-HIV; anti-inflammatory; anaphylaxis; antineoplastic; cardiac;  
E neuroprotective; valvular; gene therapy.  
X Homo sapiens.  
X

Query Match: 61.9%; Score 3373; DP 21; Length 630;  
Best Local Similarity 99.8%; Pred. No. 3.8e-284;  
Matches 628; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 17 GAGPPESEPSISRGNYTKQYVFVGHKPGRNTTQPHRLDQYIMMKGTLVIARCH 75  
Db 1 GSGPPESEPSISRGNYTKQYVFVGHKPGRNTTQPHRLDQYIMMKGTLVIARCH 60  
Qy 77 YTVYDITSHTEECYCKKLTWKSRQADVTCTRMKGKHKBCHNFIKVLKKNDALFVCG 136  
Ub 61 YTVYDITSHTEECYCKKLTWKSRQADVTCTRMKGKHKBCHNFIKVLKKNDALFVCG 120  
Qy 137 TNAPNSCRNYKMDTEPFGEESGMARCPYDAKHANVALFADGKLYSATVTOFLADAV 136  
Db 121 TNAPNSCRNYKMDTEPFGEESGMARCPYDAKHANVALFADGKLYSATVTOFLADAV 180  
Qy 197 IYSSLSGESPLRTVKHDSKWKEPYTVQAVDYSDYIFFPREIAVENYTKGVVFFRVAQ 256  
Db 161 IYSSLSGESPLRTVKHDSKWKEPYTVQAVDYSDYIFFPREIAVENYTKGVVFFRVAQ 240  
Qy 257 VCKNKGSGSRVLEKQWTSFLKARLNCSPGDSHFYFNILCAVTDVIRINGROWLATFS 316

241 VGNMGGGCRVEKQWTSFLKARVNSVPSGSHRYFNILQNTVTRINERDVVLAFTS 300  
 317 TPYNS:PGSAVCAYDMLD:ASVFTGRFEQKSPDSTWTPVDERVKEKRPCCAGSSSE 376  
 301 TPYNSIPGSAVCAYDMLDIASVFTGRFEQKSPDSTWTPVDERVKEKRPCCAGSSSE 360  
 377 RVATSEFPDPT:NIKTHPLMDA:VPSIENPWP:RTWRYELTKIADT:AAAPYQHT 436  
 361 RVATSEFPDPT:NIKTHPLMDA:VPSIENPWP:RTWRYELTKIADT:AAAPYQHT 420  
 437 VYFLGSEKGI:ILKFLARIGNSGFLNDSLFLEMSYVNSEKSYDGVDEKRWGMQLDRAS 496  
 421 VYFLGSEKGI:ILKFLARIGNSGFLNDSLFLEMSYVNSEKSYDGVDEKRWGMQLDRAS 480  
 497 SSVYVAFSTCVIKVPLGRCERHGKCKKTCIAGRPYCGWIKEGGACSHLSPNSLTFFQC 556  
 481 SSVYVAFSTCVIKVPLGRCERHGKCKKTCIAGRPYCGWIKEGGACSHLSPNSLTFFQC 540  
 557 IERKNTDGLGDCHNSFVALNGHSSLSLSTTTSDSTAGEGYESRGYKLDWKHLSDSPST 616  
 541 IERKNTDGLGDCHNSFVALNGHSSLSLSTTTSDSTAGEGYESRGYKLDWKHLSDSPST 600  
 617 DPLGAVSSHNNHCKKKGVPRESYGRGHQDL 645  
 601 DPLGAVSSHNNHCKKKGVPRESYGRGHQDL 629  
 RESULT 8  
 AAB95317  
 O AAB95317 standard; Protein; 574 AA.  
 C AAB95317;  
 X 26-JUN-2001 (first entry;  
 X Human protein sequence SEQ ID NO:17566.  
 X Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 X Homo sapiens.  
 X EP:074617-A2.  
 X 07-FEB-2001.  
 X 28-JUL-2000; 2000EP-0116126.  
 X 29-JUL-1999; 99JP-0248036.  
 X 27-AUG-1999; 99JP-0302953.  
 X 11-JAN-2000; 2000JP-0118776.  
 X 02-MAY-2000; 2000JP-0183767.  
 X 09-JUN-2000; 2000JP-0241899.  
 X (HEX-1) HEX1X RES INST.  
 X Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 X Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 X MPI; 2001-318749/34.  
 X Primer sets for synthesizing polynucleotides, particularly the 5602  
 X full-length cDNAs defined in the specification, and for the detection  
 X and/or diagnosis of the abnormality of the proteins encoded by the  
 X full-length cDNAs -  
 X Claim 8; SEQ ID 17568; 2537pp + CD ROM; English.  
 X The present invention describes primer sets for synthesising 5602  
 X full-length cDNAs defined in the specification. Where a primer set  
 X comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
 X to the complementary strand of a polynucleotide which comprises one of  
 X the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesising polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAB93166 to AAB13628 and  
 CC AAB13633 to AAB18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAB13629 to AAB13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SO Sequence 574 AA;

Query Match 54.7%; Score 2980.5; DB 22; Length 574;  
 Best Local Similarity 90.7%; Pred. No. 4.9e-250;  
 Matches 574; Conservative 0; Mismatches 0; Indels 59; Gaps 1;  
 QY 398 MDEAVDSIENRPKELSTMYPYELTKIADT:AAAPYQHT:VWF:GSEKGIILKELARIGNS 457  
 DB 1 MDEAVDSIENRPKELSTMYPYELTKIADT:AAAPYQHT:VWF:GSEKGIILKELARIGNS 457  
 QY 458 GFJNDSLFLEMSYVNSEKSYDGVDEKRWGMQLDRASSLSYVAFSTCVIKVPLGRCER 517  
 DB 21 -----CSYDGVDEKRWGMQLDRASSLSYVAFSTCVIKVPLGRCER 61  
 QY 518 HGCKKTC:ASRPYCGWIKEGGACSHLSPNSLTFFQDIERTGNTDGLGDCHNSFVALNG 577  
 DB 62 HGCKKTC:ASRPYCGWIKEGGACSHLSPNSLTFFQDIERTGNTDGLGDCHNSFVALNG 521  
 QY 578 HSSLSLSPSTTTSDSTAGEGYESRGYKLDWKHLSDSPSTCPLGAVSSHNNHCKKKGVIRES 637  
 DB 122 HSSLSLSPSTTTSDSTAGEGYESRGYKLDWKHLSDSPSTCPLGAVSSHNNHCKKKGVIRES 581  
 QY 638 YLKGHDCLPVTLIA:AVILAFVGVAFVSGITVYVCDHRRKDVAVVQKKEKELTHSRRG 697  
 DB 182 YLKGHDCLPVTLIA:AVILAFVGVAFVSGITVYVCDHRRKDVAVVQKKEKELTHSRRG 241  
 QY 698 SMSVSVTKLSGLPGDTQSKDPKPEALITPLMNSKGLATPGNTAKMLIKADCHHLDLTALPT 757  
 DB 242 SMSVSVTKLSGLPGDTQSKDPKPEALITPLMNSKGLATPGNTAKMLIKADCHHLDLTALPT 381  
 QY 756 PESTPTLQQRKFSRSRERWERNQNLINACTKDPVPMGSPVITDLPGRASPHIISVVV 817  
 DB 302 PESTPTLQQRKFSRSRERWERNQNLINACTKDPVPMGSPVITDLPGRASPHIISVVV 361  
 QY 818 LFTTQGGYQHEVYDQPKYSEVACVALEDQAATLEYKTIKEHLSSKSPNHGWNLEKLSL 877  
 DB 362 LFTTQGGYQHEVYDQPKYSEVACVALEDQAATLEYKTIKEHLSSKSPNHGWNLEKLSL 481  
 QY 878 PPKYVQREASLGPPGASLSQTL:SKLJEVHHSSSYGVYKRSYPTNSLTRSHOATLKEN 937  
 DB 422 PPKYVQREASLGPPGASLSQTL:SKLJEVHHSSSYGVYKRSYPTNSLTRSHOATLKEN 481  
 QY 936 NTSSKSSSHLSRNQSGRGNPPAPRCVDS:DVHSSQFSGQAVTVSRPSLNAYKSLTR 997  
 DB 482 NTSSKSSSHLSRNQSGRGNPPAPRCVDS:DVHSSQFSGQAVTVSRPSLNAYKSLTR 541  
 QY 998 SGLKRTPSLXPDVPPKPSFAPLST:SMKPNDACT 1030  
 DB 542 SGLKRTPSLXPDVPPKPSFAPLST:SMKPNDACT 574

RESULT 9  
 AAB94104  
 IC AAB94104 standard; Protein; 562 AA.  
 XX

AA94104;  
26-JUN-2001 (first entry)  
Human protein sequence SEQ ID NO:14328.  
Human: primer; detection; diagnosis; antisense therapy; gene therapy.  
Homo sapiens.  
EP1074617-A2.  
07-FEB-2001.  
28-JUL-2000; 2000EP-0116126.  
29-JUL-1999; 95JP-0248036.  
27-AUG-1999; 95JP-0300253.  
11-JAN-2000; 2000JP-0118776.  
02-MAY-2000; 2000CP-0183767.  
09-JUN-2000; 2000CP-0241895.  
(HELI-) HELIX RES INST.  
Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto C,  
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
WPI; 2501-318749/34.  
Primer sets for synthesizing polynucleotides, particularly the 5602  
full-length cDNAs defined in the specification, and for the detection  
and/or diagnosis of the abnormality of the proteins encoded by the  
full-length cDNAs.  
Claim 8; SEQ ID 14326; 2537pp + CD ROM; English.  
The present invention describes primer sets for synthesizing 5602  
full-length cDNAs defined in the specification, where a primer set  
comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
to the complementary strand of a polynucleotide which comprises one of  
the 5602 nucleotide sequences defined in the specification, where the  
oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
of an oligonucleotide comprising a sequence complementary to the  
complementary strand of a polynucleotide which comprises a 5'-end  
sequence and an oligonucleotide comprising a sequence complementary to a  
polynucleotide which comprises a 3'-end sequence, where the  
oligonucleotide comprises at least 15 nucleotides and the combination of  
the 5'-end sequence/3'-end sequence is selected from those defined in  
the specification. The primer sets can be used in antisense therapy and  
in gene therapy. The primers are useful for synthesizing polynucleotides,  
particularly full-length cDNAs. The primers are also useful for the  
detection and/or diagnosis of the abnormality of the proteins encoded by  
the full-length cDNAs. The primers allow chaining of the full-length  
cDNAs easily without any specialised methods. AAH3166 to AAH3628 and  
AAH3633 to AAH18742 represent human cDNA sequences; AAH32446 to  
AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
represent oligonucleotides, all of which are used in the exemplification  
of the present invention.

Query Match 54.2%; Score 2952; DB 22; Length 562;  
Best Local Similarity 100.0%; Freq. No. 1,40-247;  
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0.  
469 MSVYASEKSYGVVEDKRMVQQLDRASSLLVAFSTCVKYPAGCRPHKQKTCAS 525  
1 MSVYASEKSYGVVEDKRMVQQLDRASSLLVAFSTCVKYPAGCRPHKQKTCAS 50  
529 RPYGQWIKKEGACSHLSNSRLTFEQDTERGNTDGLDCNHSFVALNGHSSLLPSTT 586  
61 RPYGQWIKKEGACSHLSNSRLTFEQDTERGNTDGLDCNHSFVALNGHSSLLPSTT 120

589 SDSTAQEGYESRGMLDKWHLSDSPDSTPLGAVSSHHQDKKGVRESYKLGHDQLPVP 648  
121 SDSTAQEGYESRGMLDKWHLSDSPDSTPLGAVSSHHQDKKGVRESYKLGHDQLPVP 160  
649 TLLAIVILAFVAGVAFSGITVYVCDHRKQKAVVQVQKEKETHSPRGSMSVTKLSGL 708  
181 TLLAIVILAFVAGVAFSGITVYVCDHRKQKAVVQVQKEKETHSPRGSMSVTKLSGL 240  
709 FGDTQSKPKPEAILTPAXHNGKATPGNTAKYLIKADQHHLDLTALPTPESTPTLOQKR 768  
241 FGDTQSKPKPEAILTPAXHNGKATPGNTAKYLIKADQHHLDLTALPTPESTPTLOQKR 300  
769 KPSRGSREMERONLINACTKMPYKSGSPVPTDLPURASPSHIPSVAVLPTIQQYQHE 828  
301 KPSRGSREMERONLINACTKMPYKSGSPVPTDLPURASPSHIPSVAVLPTIQQYQHE 360  
929 YVQPKMSEVAQMALEDQAATLEYKTKHELSKSGSPHGVNLYENLDSLPKYPQREASL 888  
361 YVQPKMSEVAQMALEDQAATLEYKTKHELSKSGSPHGVNLYENLDSLPKYPQREASL 420  
889 GPTGASLSOTGLSKRLMEHSHSSYGVYKESYPTNSLTRSHQATTIKRNYNTSSNSHLS 948  
421 GPTGASLSOTGLSKRLMEHSHSSYGVYKESYPTNSLTRSHQATTIKRNYNTSSNSHLS 480  
949 RNOSFGRGDNPPAPQGVSTQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKATPSLKP 1008  
481 RNOSFGRGDNPPAPQGVSTQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKATPSLKP 540  
1009 DVPPKPSFAPLSTSKXPNDACT 1030  
541 DVPPKPSFAPLSTSKXPNDACT 562  
RESULT 10  
AAH93444  
ID AAH93444 standard; Protein; 562 AA.  
XX  
AC AAH93444;  
XX  
DT 06-NOV-2001 (first entry)  
XX Human polypeptide, SEQ ID NO: 3088.  
XX Human; full length cDNA; cDNA synthesis; oligo-capping.  
XX Homo sapiens.  
XX  
XX EP:130094-A2.  
XX 05-SEP-2001.  
XX 07-JUL-2000; 2000EP-0114389.  
XX 08-JUL-1999; 99CP-0194486.  
XX 11-JAN-2000; 2000CP-019774.  
XX 02-MAY-2000; 2000CP-0183765.  
XX  
XX (HELI-) HELIX RES INST.  
XX  
XX Ota T, Nishikawa T, Isegai T, Hayashi K, Ishii S, Kawai Y;  
XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX WPI; 2001-524255/58.  
XX N PSDS; AAK94365.  
XX  
XX #30 Primers useful for synthesizing full length cDNA clones and their  
XX use in genetic manipulation.  
XX  
XX Claim 8; SEQ ID NO 3088; 1380pp + sequence listing; English.  
XX The invention relates to primers for synthesizing full length cDNA  
XX clones, 830 cDNA molecules encoding a human protein have been  
XX isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA

molecules have been determined. Primers for synthesising the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the cDNA-rafting method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a polypeptide encoded by a full length human cDNA of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO.

```

Sequence      562 AA;
Query Match   54.1%; Score 2947; DB 22; Length 562;
Best Local Similarity 99.8%; Pred. No.3.8e-247;
Matches 561; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

469 MSVYNSEKCSYGVGDEKPKIMGVQIDRASSLSVYAFSCVIVKPLGRGERRGKCKKCTTAS 528
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1 MSVYNSEKCSYGVGDEKPKIMGVQIDRASSLSVYAFSCVIVKPLGRGERRGKCKKCTTAS 50
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

529 RDPYCGWIKKEGACSHLSPNSRLTFEQDIERGNTDGLCCHNSFVALNGHSSSLPSTTT 598
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
61 RDPYCGWIKKEGACSHLSPNSRLTFEQDIERGNTDGLCCHNSFVALNGHSSSLPSTTT 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

589 SCSIAQSGYSESRGGMLEKHLDDPDSTDFLGAVSSNHQCKKGVIRESYLKRGDQIVTV 648
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
121 SCSIAQSGYSESRGGMLEKHLDDPDSTDFLGAVSSNHQCKKGVIRESYLKRGDQIVTV 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

649 TLLAAVILAFVGNVSGITGVYVCDHRRKDVAVCKKEKELTSHSRGSMSSVTKLSGL 706
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
181 TLLAAVILAFVGNVSGITGVYVCDHRRKDVAVCKKEKELTSHSRGSMSSVTKLSGL 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

709 PGDTQSKDPKPEAILTPVHNKGKATPNTAKMGLKADQHILDTALTPESTETTLQCKR 768
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
241 PGDTQSKDPKPEAILTPVHNKGKATPNTAKMGLKADQHILDTALTPESTETTLQCKR 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

769 KPSRGSREWBENQNLINACTKONGPMASPVPTDLPKASPSHPSNVPLTCGGVQCHE 828
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
301 KPSRGSREWBENQNLINACTKDMFNGSFVPTDLPKASPSHPSNVPLTCGGVQCHE 360
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

929 YVDCPKMSEVAQMALEDCAAFLFKTKHEHLSKSPNHEGVNENLDSLPKVPQREASL 988
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
361 YVDCPKMSEVAQMALEDCAAFLFKTKHEHLSKSPNHEGVNENLDSLPKVPQREASL 420
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

989 GPPKASLSCTGLSRPLRMEHSSSTGVGVYKSYPTNSLTPSHQATTLKRYNTNSNSNHL 1048
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
421 GPPKASLSCTGLSRPLRMEHSSSTGVGVYKSYPTNSLTPSHQATTLKRYNTNSNSNHL 480
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

949 RNQSFEGGNGPPPAQFQVDSFTQVHSSSFSGQAVTVSQPSLMAVNSLTPSGLKRTPELKP 1009
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
481 RNQSFEGGNGPPPAQFQVDSFTQVHSSSFSGQAVTVSQPSLMAVNSLTPSGLKRTPELKP 540
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

1009 DVBPKPSTFABLSTSNKFNDACT 1030
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
541 DVBPKPSTFABLSTSNKFNDACT 562
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

RESULTS

990505

ABG04065 standard; Protein; 863 kDa.

20040666.

ABG 04255;

02-888-67

Novel human diagnostic protein. #4357.

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

human; chromosome mapping; gene mapping; gene therapy; forensic food supplement; medical imaging; diagnostic; cancer; disorder

0  
3  
0  
4  
2  
2  
1  
0  
3  
0  
0  
1

*Homo sapiens.*

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

WO200175067-A2.

11-OCT-2001.  
30-MAR-2001; 200WC-US0863..  
31-MAR-2000; 200OUS-US40217.  
23-AUG-2000; 200OUS-3649157.  
TRYSE.; HYSEQ INC.  
Dmanac RT, Liu C, Tang YF;  
WZ.; 2001-539362/73.  
NPSDB; AAS68453.  
New isolated polynucleotide and encoded polypeptides, useful in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and to assess  
biodiversity -  
Claim 20; SEQ ID No 34425; 103pp; English.  
The invention relates to isolated polynucleotide (I) and  
polypeptide (II) sequences. (I) is useful as hybridisation probes,  
polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
and gene mapping, and in recombinant production of (II). The  
polynucleotides are also used in diagnostics as expressed sequence tags  
for identifying expressed genes. (I) is useful in gene therapy techniques  
to restore normal activity of (II) or to treat disease states involving  
(II). (II) is useful for generating antibodies against it, detecting or  
quantitating a polypeptide in tissue, as molecular weight markers and as  
a food supplement. (II) and its binding partners are useful in medical  
tracing of sites expressing (II). (I) and (II) are useful for treating  
disorders involving aberrant protein expression or biological activity.  
The polypeptide and polynucleotide sequences have applications in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits to assess biodiversity  
and to produce other types of data and products dependent on DNA and  
amino acid sequences. A530001-A530337 represent novel human  
diagnostic amino acid sequences of the invention.  
Note: the sequence data for this patent did not appear in the printed  
specification, but was obtained in electronic format directly from WIPO  
at <http://wipo.int/pub/published> not sequences.

```

Query Match      51.0%   Score 2891:   DB 22:   Length 863:
Best Local Similarity 80.9%   Pred. No. 5.9e-42:
Matches 583:   Conservative 23:   Mismatches 92:   Indels 18:   Gaps 8

QY      321  S P G S A V A C A Y L I A I A S V F T R F E K Q S P S T W T P V D E R V P R P G C C A G S S L E R V A T 360
DB      150  S P G S A V A C A Y L I A I A S V F T R F E K Q S P S T W T P V D E R V P R P G C C A G S S L E R V A T 200
QY      381  S N E F P C D T L N I K T H P L M D E A V P S I F R P R F L A T M V Y R L T K I A V T A A G P Q N H T W P F 440
DB      210  S N E F P C D T L N I K T H P L M D E A V P S I F R P R F L A T M V Y R L T K I A V T A A G P Q N H T W P F 269
QY      441  G S E K I I L K P A R T O N S O F L A N S I F J E M S V Y N S E K S Y D G V E D K I - M M Q L D R A S S L 499
DB      270  G S E K I I L K P A R I G N S O F L A N S I F J E M S V Y N A S E K - K N S T A K P V R V T I I L N P G Q A S F C I 328
QY      500  Y V A F S T C V - - - - K V P - G R C - E R H G K C K T C I A S R D P V C - - - G W I K E G G A C S H L S P N S R 550
DB      329  T I R E V C V A R R K H - K C P P Y R C T L Q V H - F C P C H L S G K E T I C R V T G G M K V A A D R D E S - P Y A A 367
QY      581  L P F E C C I - E R O N T D G L G C H N S P V A L N G H S S L S P S T T T S D S T A Q C E Y S E R S G M L D W K H L 609
DB      388  M L A A C M A C R K E L G I T A L H I K H R A T G N R T K - - - - T P G P G A S R S S F C P L G C L K W Q L 442
QY      610  L D S P D S T D P L G A V S S H N C Q K K V I R S S Y - K G H D Q L V P V T L L A A V I L A F V M G A V F S G I T 669
DB      443  F P R R L R M P O G S R R K R S Q L E A Q R V I R S S Y L K G H D Q L V P V T L L A A V I L A F V M G A V F S G I T 502

```

670 VYVCDHRRKDAVYVORKEKELTISRRGYSVSVTK:SGLEGTQSKDKPKGPAITPLMHN 729  
|||||  
503 VYVCDHRRKDAVYVORKEKELTISRRGYSVSVTK:SGLEGTQSKDKPKGPAITPLMHN 562  
|||||  
710 GKLATPGNTAKMLIKACCHHLDJALPTPESTPT:COKEKTSRGRERERNQNLINACTK 789  
|||||  
563 GKLATPGNTAKMLIKACCHHLDJALPTPESTPT:COKEKTSRGRERERNQNLINACTK 522  
|||||  
790 KNPFGMSVPIPTDLRLRARDPSHIFSVAV:PIQQGYQYQYVYVQPKYSEVACMALEDQAT 849  
|||||  
623 EYPMGSPV:PTDLRLRARDPSHIFSVAV:PIQQGYQYQYVYVQPKYSEVACMALEDQAT 582  
|||||  
950 LEYKTIKEHLSSKSPNHNQVJENLQSLPPKVPQREAS:GPPGASLSQTGSKRLMEHHS 909  
|||||  
683 LEYKTIKEHLSSKSPNHNQVJENLQSLPPKVPQREAS:GPPGASLSQTGSKRLMEHHS 742  
|||||  
910 SSYGYEYKASYPTNSLTRSHQATILKRNNTNSNSASH:SPNQSFGRGNPPAPQVLSI 969  
|||||  
743 SSYGYEYKASYPTNSLTRSHQATILKRNNTNSNSASH:SPNQSFGRGNPPAPQVLSI 802  
|||||  
970 QVHSSQPSQAVTVSRQPELKNAYNSLTRSGUKRTSLKFDVDPPEKSPAPLSTSKYKNNAC 1029  
|||||  
923 QVHSSQPSQAVTVSRQPELKNAYNSLTRSGUKRTSLKFDVDPPEKSPAPLSTSKYKNNAC 862  
|||||  
1030 T 1030  
863 T 863  
RESULT 12  
ABG04620 standard; Protein; 962 AA.  
ABG04620;  
13-FEB-2002 (first entry)  
Novel human diagnostic protein #4611.  
Human; chromosome mapping; gene mapping; gene therapy; forensic;  
food supplement; medical imaging; diagnostic; genetic disorder.  
Homo sapiens.  
WO250175067-A2.  
11-OCT-2001.  
3C-MAR-2001; 2001WO-US05631.  
31-MAR-2000; 2000US-0542217.  
23-AUG-2000; 2000US-0649167.  
(HYSE-) HYSEQ INC.  
Drmanac RT, Liu C, Tang YT;  
WPI; 2001-039162/73.  
N-PSDB; AAS68807.  
New isolated polynucleotide and encoded polypeptides, useful in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and to assess  
biodiversity.  
Claim 20; SEQ ID No 14979; 103pp; English.  
The invention relates to isolated polynucleotide (I) and  
polypeptide (II) sequences. (I) is useful as hybridisation probes,  
diagnostics, forensics, gene mapping, oligomers, and for chromosome  
polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
and gene mapping, and in recombinant production of (II). The  
polynucleotides are also used in diagnostics as expressed sequence tags

for identifying expressed genes. (I) is useful in gene therapy techniques  
to restore normal activity of (II) or to treat disease states involving  
(II). (II) is useful for generating antibodies against it, detecting or  
quantitating a polypeptide in tissue, as molecular weight markers and as  
a food supplement. (II) and its binding partners are useful in medical  
imaging of sites expressing (II). (I) and (II) are useful for treating  
disorders involving aberrant protein expression or biological activity.  
The polypeptide and polynucleotide sequences have applications in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits to assess biodiversity  
and to produce other types of data and products dependent on DNA and  
amino acid sequences. ABG04620-A5G10377 represent novel human  
diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
specification, but was obtained in electronic format directly from Wipo  
at ftp.wipo.int/pub/published\_pat\_sequences.

Sequence 962 AA;  
Query Match 52.88; Score 2879.5; DB 22; Length 962;  
Post-Local Similarity 70.33; Pred. No. 7.1e-241;  
Matches 593; Conservative 30; Mismatches 122; Gaps 12;  
147 YKVDY-LEPFGDEFSGMARCPYDAKHANVALFA-----DGKLYSATVTCDFLAD 194  
|||  
95 YEPDTWLGKWSSTWLVFTHPFFSAALTHSALTARSUTGSLTSPDGKLYSATVTCDFLAD 154  
|||  
195 AV-YRSLGESPTLRTYKHSKWKKEPYFQAVQYGDY-YFPFREIAVEYNTYKVVPRV 254  
|||  
155 AV-YRSLGESPTLRTYKHSKWKKEPYFQAVQYGDY-YFPFREIAVEYNTYKVVPRV 214  
|||  
255 AQVCKNDMGSGQVLEKQWTSF-KARLNCVSPGDSHFYFNILQAVTDVIPINGRVYLAT 314  
|||  
215 AQVCKNDMGSGQVLEKQWTSF-KARLNCVSPGDSHFYFNILQAVTDVIPINGRVYLAT 274  
|||  
315 PSTPYNSIPGSAVCAVEXLDIASVFTGRFKEKSPDSTWTPVDERVAPKPRPCCAGSSS 374  
|||  
275 PSTPYNSIPGSAVCAVEXLDIASVFTGRFKEKSPDSTWTPVDERVAPKPRPCCAGSSS 334  
|||  
375 LERYATSNPEPDJLNFIKTHPLMDSAVESIFNRPWFLRTVRYRJTKIADVTAAAGPYQN 434  
|||  
335 LERYATSNPEPDJLNFIKTHPLMDSAVESIFNRPWFLRTVRYRJTKIADVTAAAGPYQN 394  
|||  
435 HTVYFLGSEKGIILKFLARIGNSGFNDSLFLBEMSVYNSEKSYDGVEDKRMGVQDLR 494  
|||  
395 HTVYFLGSEKGIILKFLARIGNSGFNDSLFLBEMSVYNSEKSYDGVEDKRMGVQDLR 454  
|||  
495 ASSSLYVAFSTCVIKVPLGRCERHGHCKKTCIASRDPYCGWIKEGGACSHLSPNSRLT-- 552  
|||  
455 ASSSLYVAFSTCVIKVPLGRCERHGHCKKTCIASRDPYCGWIKEGGACSHLSPNSRLTORG 514  
|||  
553 FEQDIERGNTDGLGDC-----HNSFVAJNGHSSS----- 581  
|||  
515 FVWFVVTGST--LHRCILLITSLWQTHPSY-----HSSSETGRPVKXGGERGKCTLNHFEL 567  
|||  
582 --LPSTTTSCSTAQEGYESRGGM-----LWKHGLLDSFDSTDPGLGAVSSHNHCKKGVII 634  
|||  
569 GWKLFVSYTLLRVVIGVGLKRCOTLFFRRRLRW-----PQGRRRKRSQLEAQRV-- 616  
|||  
635 RESYLKRGHCDQVNTLLA-AVILAFVMGAVFSGITVYCVCDHERKDVAVVQVORKEKELTHS 694  
|||  
617 RESYLKRGHCDQVNTLLA-AVILAFVMGAVFSGITVYCVCDHRKDVAVVQVORKEKELTHS 676  
|||  
695 RRSSMSVTKLSGLFDQTSKQPKPEALITPLMHNKGLATPGNTAKMLIKACDGHLLDFTA 754  
|||  
677 RRSSMSVTKLSGLFDQTSKQPKPEALITPLMHNKGLATPGNTAKMLIKACDGHLLDFTA 736  
|||  
755 LPTPESTPLQOKRKPSRGRERERNQNLINACTKXPPMGSPVITDPLRASPSPHIPS 814  
|||  
737 LPTPESTPLQOKRKPSRGRERERNQNLINACTKXPPMGSPVITDPLRASPSPHIPS 796  
|||  
915 VVVLPIITCCGYQYQYVYVQPKYSEVACMALEDQATLEYKTIKEHLSSKSPNHNQVJENL 974  
|||

D 797 WNSCPSSRATSMSTWTPXAXRWPWRKTRTPHNSIFPSNINISAAKVTPTNEXILKRW A56  
Y 875 USLPPK-----VPCREASLGPFGAS-----LSCTLSKELNEMHSSSYGVYKRSY 920  
D 657 TACPPKPHSRPPWPVRPCPLRFVXASGCKCTPLTTLTIA----- 999  
Y 921 PINSLTRSHQAT 932  
b 900 GATPRTSRHEAT 911  
  
RESULT 13  
AB92688  
D AAB92688 standard; Protein; 507 AA.  
X  
X AAB92688:  
X 26-JUN-2001 (first entry)  
X Human protein sequence SEQ ID NO:1073.  
E Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
W Homo sapiens.  
X EPI074617-A2.  
X 07-FEB-2001.  
X 26-JUL-2000; 2000EP-016126.  
X 29-JUL-1999; 96JP-0348236.  
X 27-AUG-1999; 96JP-0300253.  
X 11-JAN-2000; 2000JP-0118776.  
X 02-MAY-2000; 2000JP-0183767.  
X 09-JUN-2000; 2000JP-0241F93.  
X (HELI-1) HELIX RES INST.  
X Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto G,  
X Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T,  
X WPI; 2001-318749/34.  
X Primer sets for synthesizing polynucleotides, particularly the 5602  
X full-length cDNAs defined in the specification, and for the detection  
X and/or diagnosis of the abnormality of the proteins encoded by the  
X full-length cDNAs -  
X Claim 8; SEQ ID 11073; 2537pp + 1D PCM; English.  
X The present invention describes primer sets for synthesizing 5602  
X full-length cDNAs defined in the specification, where a primer set  
X comprises: (a) an oligo-dT primer, and an oligonucleotide complementary  
X to the complementary strand of a polynucleotide which comprises one of  
X the 5602 nucleotide sequences defined in the specification, where the  
X oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
X of an oligonucleotide comprising a sequence complementary to the  
X complementary strand of a polynucleotide which comprises a 5'-end  
X sequence and an oligonucleotide comprising a sequence complementary to a  
X polynucleotide which comprises a 3'-end sequence, where the  
X oligonucleotide comprises at least 15 nucleotides and the combination of  
X the 5'-end sequence/3'-end sequence is selected from those defined in  
X the specification. The primer sets can be used in antisense therapy and  
X in gene therapy. The primers are useful for synthesizing polynucleotides,  
X particularly full-length cDNAs. The primers are also useful for the  
X detection and/or diagnosis of the abnormality of the proteins encoded by  
X the full-length cDNAs. The primers allow obtaining of the full-length  
X cDNAs easily without any special-used methods. AAH01166 to AAH13628 and  
X AAH13633 to AAH15742 represent human cDNA sequences; AAB92446 to  
X AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
X represent oligonucleotides, all of which are used in the exemplification  
X of the present invention.

XX Sequence 507 AA:  
SQ  
Query Match 45.1%; Score 2621.5; DB 22; Length 507;  
Best Local Similarity 90.2%; Pred. No. 7e-219;  
Matches 507; Conservative 0; Mismatches 55; Gaps 1;  
469 MSVYNSEKSYDGVEDKRINXGLDRASSSLYVAFSTCVIKVPLGRCEHKGCKKTCIAS 528  
D 1 MSVYNSEKSYDGVEDKRINXGLDRASSSLYVAFSTCVIKVPLGRCEHKGCKKTCIAS 60  
QY 529 RDPYCGWKEGGASHLSPNSRLTFEQDIBRGTDGLGCHNSFVALNGHSSLSLPTTT 588  
D 61 RDPYCGWKEGGASHLSPNSRLTFEQDIBRGTDGLGCHNSFVALN----- 108  
QY 589 SDSTAQEGYESRGMDLWKELLSPOSTDPLGAVSNNHCKKGVIRSYLKGHDQLVPV 648  
D 109 -----GVIRSYLKGHDQLVPV 125  
QY 649 TLLAIAVILAFVMAVSGITVYVCVCHRRKDVAVVQVQKKEKELTHSRGSMSSVTKLSGL 708  
D 126 TLLAIAVILAFVMAVSGITVYVCVCHRRKDVAVVQVQKKEKELTHSRGSMSSVTKLSGL 185  
QY 709 PGDTQSKDPKPEALITPLMENGKLAIPGNTAKVLIKADQHHLLDLTALPTPESTPTLQCKR 768  
D 186 PGDTQSKDPKPEALITPLMENGKLAIPGNTAKVLIKADQHHLLDLTALPTPESTPTLQCKR 245  
QY 769 KPSRGSREWERNONLINACTKDYPPMGSPVITDPLRASPSHIPSVVVLPITQQGYCHE 828  
D 246 KPSRGSREWERNONLINACTKDYPPMGSPVITDPLRASPSHIPSVVVLPITQQGYCHE 925  
QY 829 YVDPQKYSVAQVALEDCAAITLEYKTKHEHLSKSPNKGVLNVLNLSLPPKVPQREASL 888  
D 306 YVDPQKYSVAQVALEDCAAITLEYKTKHEHLSKSPNKGVLNVLNLSLPPKVPQREASL 365  
QY 889 GPPGASLSQITGLSKLEEMHSSSYGVYKSYPTNSLTPSHQATT--KRNNTNSNSSHLS 948  
D 366 GPPGASLSQITGLSKLEEMHSSSYGVYKSYPTNSLTPSHQATT--KRNNTNSNSSHLS 425  
QY 949 ENQSFGRGDNFPFPAPQVDSIQVHSSQPSGQAVTVSOPSLNAYNSLTRSGLKRTPLSLKP 1006  
D 426 ENQSFGRGDNFPFPAPQVDSIQVHSSQPSGQAVTVSOPSLNAYNSLTRSGLKRTPLSLKP 485  
QY 1009 DVPPKPSFAPJSTSMKKNDACT 1030  
D 486 DVPPKPSFAPJSTSMKKNDACT 507  
  
RESIDUE 14  
AAG63213  
ID AAG63213 standard; Protein; 1086 AA.  
XX AC AAG63213;  
XX  
DT 01-OCT-2001 (first entry)  
D2 Amino acid sequence of a human semaphorin-like polypeptide.  
XX Expressed sequence tag; EST; semaphorin-like protein; neuronal growth;  
XX spinal cord damage; neurodegenerative disease; genetic neuronal defect;  
XX immunological disorder; lymphocyte dysfunction; viral infection; cancer.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH Peptide 1..16  
FT /note= "Signal peptide"  
FT Domain 671..693  
FT /note= "transmembrane domain"  
XX WC020153466-A1.  
XX 26-JUL-2001.  
PD

22-DEC-2000; 2003KO-US33158.  
 23-DEC-1999; 99US-0471275.  
 21-JAN-2000; 2000US-0488725.  
 25-APR-2000; 2000US-0552317.  
 31-AUG-2000; 2000US-0653274.  
 (HYSF-) HYSEQ INC.  
 Boyle B., Yeung GY, Atterburn VC, Nize NK, Tang YF, Liu C.  
 Dimaac RT:  
 WPI: 2001-476114/51.  
 N-PSDB; AA42197.  
 New polynucleotides and the encoded polypeptides having human secreted semaphorin-like polypeptide, useful in treating neurodegenerative diseases, viral infections or cancers, or diagnosing and mapping genetic neuronal defects  
 Claim 12: Page 124-127; 142pp; English.  
 The present sequence represents a human semaphorin-like polynucleotide. Semaphorin-like polynucleotides and polypeptides are useful in therapeutic, diagnostic and research methods. These are particularly useful in modulating neuronal growth regenerative capacity (e.g. in the case of spinal cord damage), treating neurodegenerative diseases, diagnosing and mapping genetic neuronal defects. These are also useful in treating immunological disorders arising from T and B lymphocyte dysfunction, or in treating viral infections or cancers.  
 Sequence 1086 AA:  
 Query Match 43.1%; Score 2361.5; DB 22; Length 1086;  
 Best Local Similarity 45.5%; Pred. No. 1.1e-195;  
 Matches 512; Conservative 159; X.matches 306; Indels 147; Gaps 30;  
 1 MRSEALLVFTLL---HFAGAGFEDSDSPISISHGNYTKQYVGVGHKPGNTTORHLD 57  
 1 MRVFLLCAYLLLMVSQLRAVDFEEDCEPLNTVYHVSQVYVFRG-RPSNGSQ-HRLD 58  
 58 IQVIMMGTEVIAARDHYTYDICTSTHEEYCSKKLTWKSQADVDTCRKGKHDEC 117  
 59 PQMKIKRQTSIAGRDQYVYKLEMPKTEVPPKLTWRSQDRENCVAKGKHDEC 119  
 118 HRFIKVLLKNDLALFVOTNAPNSCRNKKYKDTLEPFGDEPSHWARCPYDAKHANALP 177  
 119 HNFIKVPRNDEKVFVCGTNAPNFCRYRSTLEYEGEEISGLARCFPDARQTNALP 178  
 178 ADGKYYSATVDFALDAVYRSLGSPTLRLVKIDSKWLKZFFVQAVDYGCYIYFFER 237  
 179 ADGKYYSATVADFLASDAVYRSMGDGSAIRTKYDSMKKEPHEFLHAEYGNYYIYFFER 238  
 238 ELAVENYTKGVYFVRAVQCKNNGSGRVEKQKTSFLKAPLNCVPGSGHVFENLIQ 297  
 239 EAVENHNLKAVYRVARCKNNGSGRVEKQKTSFLKAPLNCVPGSGHVFENLIQ 298  
 298 AVTDVIRNGRWALATPTPNSIFGSAVQAYMDLDAVFTGRFKQKSPGSPWTYTP 350  
 299 SITDIQNGITVGVFTQLNSIFGSAVQAFNMDLEKVPGRFKQKSPGSPWTYTP 359  
 358 DERVPKPRGCGAGSSLEHYATSEFPDCTINFTKPLDZEVNFSINRWFEETVVP 417  
 359 EDKVPKPRGCGAGSLADAYATSIDPDETUSEFKSHFLMDGSAVPIADPWTYTP 418  
 418 YRLTKIAYDTAGPYQNHVTVFLGSEKGLKFLASLNSGF-LNDSLEFENYSWYSEK 474  
 419 YRLTKIAYDHSAGPYQNYTVTVFVSGAGVFLVLAKE-TSPFSLENDVLEDEIAYNAK 475  
 477 CSYDGVEDHRIIMQJCRASSLSLYAFSTCVIKVPLGRCKGCKKTKCIASPPYCCWI 516  
 477 CSAENEEDKKVISLQDKDHHALYVAFSSCIIPFLSRCEYVGSCKKSCIASRDPYCCWL 516

QY 537 KEGCAASHLSPNSLIT-----FEDIERGNTDGLGDCHNSFVALNGHSSSL 582  
 DB 537 SQ-GSGGRVTPGMLJLTEDFFAFHNHSAEGYQDTEFGNTAHLGDOCH-----EI 584  
 QY 583 LPSTTTSDTSAGEQVESRGG-----MLD-WKHLA-----610  
 DB 585 LPTSTTPD-----YKIFGQPTSDMEVSSSVTTMASIPEITPKVIDTWPKTSRKFV 618  
 QY 611 --DSP---DSTPLGAVSSHHCQKKGVRRESYKGDQDLVPVTLAIAVILAFYGVAVP 665  
 DB 619 VQDQENTSDFTDPLSGI-----PKGVMEVQSGESQVYHNMVLTTCVFAAFVLGAPI 691  
 QY 666 SGITVTCVCD-HERKQVAVVORKEKELTHSRSSMSVTKLSGLFG-----DTQSKDKPKE 720  
 DB 692 AGVAVCYRDMFYKARKI--HKDAESAQCTDSSGSFAKLNGLDFSPVKYQCN:DSPX 749  
 QY 721 ALLTELHNSKILATPGNTAKVLKADQHHLDLTPALTPSTPTLQQRKPSRGSREKERN 780  
 DB 750 LYNNLJTSRKELPPNCTKSNVMDHRCQPELJALPTPESTPVLRKQTLQAMKSHSEKAH 859  
 QY 781 QNLINACTKDMPPMGSEVITDLPLEASFSHPISVVVLP:TCQGYOHEY-----929  
 DB 810 GH-GASRKETPQFFSSPP2HSPL--SHGHIPSAIVLPNATHDYNTSFNSNAKAEKK 865  
 QY 830 ---VQCP---KMSEVACMALEDQAALEYKTIYEHLSKSPN-----HGYNLVE 872  
 DB 866 LQNDHPLTLYSSKREHPRASVSRNT--NLLKHLNDPNKSNKAIMGE:QMAHOKMLD 923  
 QY 873 NLDSJ---PKVQCRFASLQPPGASLSQGLSKRLEMHSSSY---GVNYSKRYPTNLSL 926  
 DB 924 PXSMSVPPKQVNRASLYSPFSTLPRKSPTRKVDVPTTPSVPMTSLEJQRGYHKNSSC 983  
 QY 927 RSHQATTLKANTNSNSHLSRQSGFRGD-NPPAPQAVDS:QVHSSQFSCQAVTVSR 985  
 DB 964 R-HSISAMPK-NLNSPGLVLSRQPSMNRGGYPTTGAKVQVQI-----GTPSVHL 1034  
 QY 986 QPSLNAYNSLT-----RSGKXTPSLKDPVPPKPSFAPLSTMKP 1025  
 DB 1035 QPSLGRSSYTSNGT:PRTGKRTPSLKDPVPPKPSFVQTPSVRP 1086  
 RESULT 15  
 AEG79177  
 ID AEG79177 standard; Protein; 1088 AA.  
 AC AEG79177:  
 XX 15-NOV-2002 (first entry)  
 DE Human semaphorin-like protein #5.  
 DE Human; NCNV; neurodegenerative disease; Alzheimer's disease; anxiety;  
 KW Parkinson's disease; Huntington's disease; neurological disorder;  
 KW schizophrenia; manic depression; mental retardation; angina pectoris;  
 KW cardiovascular disease; acute heart failure; myocardial infarction;  
 KW muscular disease; muscular disorder; retinal disease; photoreception;  
 KW gastritis; keratinisation disorder; cancer; ovarian cancer; melanoma;  
 KW immunological disorder; inflammatory disease; immune disease; diabetes;  
 KW bacterial infection; fungal infection; protozoal infection; obesity;  
 KW viral infection; reproductive system disorder; metabolic disturbance;  
 KW anorexia; wasting disorder; chronic disease; infectious disease;  
 KW dyslipidaemia; cub; sushi; myelin; von willebrand factor; kielin;  
 KW semaphorin; serine/threonine protein kinase; TGF-beta binding;  
 KW mas proto-oncogene; ribonuclease pancreatic precursor; aminotransferase;  
 KW colloid-like 2; cysteine sulfinic acid decarboxylase.  
 OS Homo sapiens.  
 PK WO200264791-A2.  
 PD 22-AUG-2002.  
 XX



10-DEC-2001; 2001WO-US48369.

08-DEC-2000; 2000US-254329P.

14-DEC-2000; 2000US-255648P.

15-MAY-2001; 2001US-29-0372P.

08-JUN-2001; 2001US-297173P.

08-JUN-2001; 2001US-309258P.

29-AUG-2001; 2001US-315639P.

01-OCT-2001; 2001US-326393P.

(CURA-) CURAGEN CORP.

Alsobrook JP, Anderson DW, Burgess CE, Boudog FL, Casman SJ,

Colman SD, Edinger SR, Eilerman K, Gerlach V, Gorman JJ, Grasse WM,

Guo X, Herrmann J, Kexuda R, Lopley DM, Li L, MacDougall JR,

Millet I, Pena CE, Peyman CA, Rastelli L, Rieger DK, Shinkate RA,

Smithson G, Spytex KA, Stone MJ, Tchertev VT, Vernet CA, Voss E2,

Zerhusen SD, Zhong H, Zhong M;

WFI; 2002-643486/63.

N-PSDS; ABS64384.

New NOVX polypeptides and polynucleotides useful for treating or

preventing e.g. neurodegenerative diseases, neurological disorders,

cardiovascular diseases, muscular diseases and disorders, or

immunological diseases

Claim 1; Page 56; 299pp; English.

The present invention relates to new NOVX polypeptides. The polypeptides,

polynucleotides and antibodies are useful in the manufacture of a

medicament for treating or preventing neurodegenerative diseases (e.g.

Alzheimer's disease, Parkinson's disease, or Huntington's disease),

neurological disorders (e.g. anxiety, schizophrenia, manic depression or

mental retardation), cardiovascular disease (e.g. acute heart failure,

angina pectoris or myocardial infarction), muscular diseases and

disorders, retinal diseases (including those involving photoreception,

deafness and keratinisation disorders), cancer (e.g. ovarian cancer or

melanoma), immunological disorders, inflammatory and immune diseases,

bacterial, fungal, protozoal and viral infections, and reproductive

system disorders. The proteins of the invention may be used to screen

drugs or compounds that modulate the NOVX protein activity or expression,

as well as to treat disorders characterised by insufficient or excessive

production of NOVX protein or protein forms that have decreased or

aberrant activity compared to NOVX wild type protein, such as diabetes,

obesity, metabolic disturbances associated with obesity, anorexia and

wasting disorders associated with chronic diseases and various cancers,

infectious diseases and various dyslipidaemias. The nucleic acid

sequences of the invention may be used in chromosome mapping.

C identifying an individual from a pure biological sample (tissue typing),

C and in forensic identification of a biological sample. The present

C amino acid sequence represents a NOVX protein of the invention.

X Q Sequence 1028 AA;

Query Match 43.1%; Score 2349.5; DB 23; Length 1068;

Best Local Similarity 45.4%; Pred. No. 122e-154;

Matches 512; Conservative 156; Mismatches 309; Indels 149; Gaps 31;

Y 1 MRSEALLLYFELL---HFAGAGPFEDSEPSI---SHGNYTKQYVFGVGHKPGNTTQPHR 45

b 1 MRVLLCAYILLMVLSQLRAVSFEDDEPLNTVDYHCKSSRQYVFRG-RFSKNEQ-HR 98

Y 56 LDQVIMIMKGLIYIARDHIYTDITSTHEEYCSKLTWKSQADVDTCNKQKHKD 115

b 59 LDFQLMLKIRBTLVIAGRDQVYTNLNEPKTEYVWQKLTWESRQCDRENCAYKQKHKD 118

Y 116 ECHNFIKVLKKNDDA-FVGTGNFNSCRWYKVDLTPFGDSFGMARCPYAKANVA 175

b 119 ECHNFIKVFVRNDEMVFVCTGNFNPYRYRYSTLVEYDGEISGLIARCPFDARQTNVA 178

Y 176 LFADGKLYSATVDFLADAVIYNLSGESPTLRTVNHDSKWLKEPFPVQAVYGVYVFF 235

179 LFADGKLYSATVDFLADAVIYNLSGESPTLRTVNHDSKWLKEPFPVQAVYGVYVFF 238

QY 236 PREAVDYNTIMKGYTFPRVAQVCKNDMGSCORVLEKQWTSFLKARLNCSPVGSHEFYFNI 295

DB 239 PREAVEHNHNLGKAVSRVARICKNDMGSCORVLEKHTWTSFLKARLNCSPVGSFYPFDV 298

QY 296 LOAVTDVIRINGRDVWJATFSTPYNSIPSSAVCAKMLDIASVETGRPRECKSPDSTWTP 355

DB 299 LOSITDIQINGFTVVGVTFTQNSIPSSAVCAFMDDIEKVFGRKECKTPDSSVNTA 358

QY 356 VPDERVPRKPRGCCAGSSSLERVATSNFEDDTLNFPIKTHPLJLDEAVPSIFENRPFWRMT 415

DB 359 VPEKVPKPRPGCCCKKGLAEAYKTSIDFDETLSPFKSHPLMDSAVPPIADPFWFKTR 418

QY 416 VRYLTIAVDYTAAGPYQNHVTVFLGSEKGIILFLARIGNSPF-LNDSJLEEMSYNS 474

DB 419 VRYLTAISVDHSAGPYQNTVTFVSGAGWILKVLAK--TSPPSLNDSVLEIEAYNH 476

QY 475 EKCSYDGVVEKRIKMGMDLRASSLYVAFSTCVIKVPLGHCERHGKCKTCIASRDPYCG 534

DB 477 AKCSABNEEDKXVLSLODKDHALYVAFSSCIIRPLSRCEYRGSCCKKCIASRDPYCG 536

QY 535 WIKEGGACSHLSPNSRLT-----PEQDIERGNTDGLGCHNSFVALNCRSS 580

DB 537 WLSQ-GSGGAVTPGMULLTEDFFAFHNHSABGYEQDTEFCNTAHLGQCH----- 584

QY 581 SJLPSTTTSSTAGEYSESQG-----YKIFGFTSDMEVSSSVYTWASIFE-TPKVIDTWRPKLTSSRK 610

DB 585 ECLPSTSTPD-----DSTPLGAVSSHHCCKGVIRSYLKGHDQJQVETLLAIVILAFVWGA 638

QY 611-----DSP-----PKGVNWEVQSGESNQVHMVLLTVCFAAFVLGA 663

DB 639 FVWQDDNTSDFDPLSGI-----HKAESAQSCTDSSGSFAKINGLDFSPVKEVQQNIDS 691

QY 664 VFSGITVYVCCD-HRRKDVAVQPKELTHSRPGSMSSVTKLSGJFS-----DTQSKDKPK 718

DB 692 FIAGVAVCYRDMFVRKVKI--HKDAESAQSCTDSSGSFAKINGLDFSPVKEVQQNIDS 749

QY 719 PEALTLPMNGKLTATQNTAKMLIKADQHHLDLTALPTPESTFLQOKKPKSGSREWE 778

DB 750 PKLYSNLTLTSKELPPKGDTSKVMYMDHRRGPPPELAALPTPESTFVLHQKTLQAMKSHSEK 809

QY 779 RNQNLINACTKQVPMGSPVPIPTDLPURASPSHPSVWVLPITCCQGYQHEV----- 829

DB 810 AFQH--GASRKEETCFPFSSPPPHSPL--SHGHIPSAIVLPLNATHDNTSFSNNAKAE 865

QY 830-----VQCP---KMEFVAQWALDDQATLETKTIKEHLSSKSPN-----HGVNL 870

DB 866 KKLQMLDHLPLTKSSKXKDHRRSVDNRNTL--NDLLKHLNDPNSNPKAIMEDIQWAFQNLN 923

QY 871 VENLDGL---PFKYPQREASLGPFGASLSCTGLSKRLEMHSSSY---GVYKESYFINS 924

DB 924 LCPMGSSSEVPKYPNREASLYSPSP-LPKNSPTKRVIVTTPGVMTSJERGQYGHNS 963

QY 925 LTRSHQATTLAKNNITKSNSSHLSRNSQSGRGC-NPQPAFCRVDSICWHSQSPSGCAVTV 963

DB 984 SCR-HS-SAMKP-NLNSPQVLLSRQPSMNRGGYMPPTGAKVDYQ-----GTPSV 1014

QY 984 SRQPSLNAYNSLT-----ESGLKETPSLKPDPVPPKPSFAPLSTSMKP 1025

DB 1035 HLQPSLSRQSSYTGNTLPRGLKSTPSLKPDPVPPKPSFAPLSTSMKP 1082

Search completed: October 23, 2003, 17:10:34

Job time : 56 secs

1	5422.5	99.5	1047	11	US-C9-957-187-85	Sequence 95, App1
2	492.	90.3	939	11	US-C9-957-187-4	Sequence 4, App1
3		84.2	824	11	US-C9-957-187-6	Sequence 6, App1
4	4532.5	94.2	824	11	US-C9-957-187-6	Sequence 6, App1
5	3373	61.9	610	11	US-C9-957-187-30	Sequence 30, App1
6	3362	61.7	626	11	US-C9-957-187-35	Sequence 35, App1
7		36.9	888	12	US-10-140-427-544	Sequence 54, App1
8	2033.5	36.9	888	12	US-10-140-427-544	Sequence 54, App1
9	2009.5	36.9	888	12	US-C9-931-826-35	Sequence 35, App1
10	2009.5	36.9	888	12	US-C9-931-826-35	Sequence 35, App1
11	2009.5	36.9	888	12	US-10-140-427-544	Sequence 54, App1
12	2009.5	36.9	888	12	US-10-140-427-544	Sequence 54, App1
13	2009.5	36.9	888	12	US-10-140-427-544	Sequence 54, App1
14	2009.5	36.9	888	12	US-10-140-427-544	Sequence 54, App1
15	2009.5	36.9	888	12	US-10-140-427-544	Sequence 54, App1

```

Query Match      99.3%; Score 5422.5; DB 11; Length 1047;
Eseq. Local Similarity 98.3%; Pred. NO. 0;
Matches 1029; Conservative 0; Mismatches 17; Gaps 1
QV      1  MRSEALLVYTLTHFAGAGPPDSEI-SISHGNYTKQYVFVGHFGKGMTTORHRLDIQM 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DQ      1  MRSEALLVYTLTHFAGAGPPDSEI-SISHGNYTKQYVFVGHFGKGMTTORHRLDIQM 60
QV      61  IMIKNGTNTYAMRCHIVYVD-ITSH-EEIYCSKSLTKWRCADVETCRMKGSHKDECNF 120

```

```
61 INXMGITVLAARDHITVVIDTSHTEETVCSKLTWKSACADVDTCRMKGKGCREF 120
121 IKVLLKXCDALFYCGTNAFPCENFKNMOTLEFFDFSSQVACQYDAKHANVALFAG 140
121 IKVLLKXCDALFYCGTNAFPCENFKNMOTLEFFDFSSQVACQYDAKHANVALFAG 140
181 KLYSATVTDLAIDAVIYRS:GESPTLRTVAKHDKWKEKPYFQAVDGYIYFFREIA 240
181 KLYSATVTDLAIDAVIYRS:GESPTLRTVAKHDKWKEKPYFQAVDGYIYFFREIA 240
241 VEYTMGVVPPVAVQVCKDMGSOQVLEKQNTSF:KAPLNGSVPGDSHFVYNIQAVT 300
241 VEYTMGVVPPVAVQVCKDMGSOQVLEKQNTSF:KAPLNGSVPGDSHFVYNIQAVT 300
301 DVIRINGRDVVLAFTSTPYRS:IPCSAVCAVDMLDIASVFTGRPEQKSPSTWTFV 360
361 VPKRPPCCAGSSSLERYATSNFDDTLAFIKTHPLMDEAVPS:FNRPFLSTWYAYL 420
361 VPKRPPCCAGSSSLERYATSNFDDTLAFIKTHPLMDEAVPS:FNRPFLSTWYAYL 420
421 TKAVDTAAGPYQNTVVFILGSEKGIILKPLARIGNSGF:NDSLFLEKSVYNSEKSYD 480
421 TKAVDTAAGPYQNTVVFILGSEKGIILKPLARIGNSGF:NDSLFLEKSVYNSEKSYD 480
481 GVEDKRTMGXQLDRASSSLVAFSTCVIKVPLGRCEHKGKCKTC:ASRDPYCGWKEGG 540
481 GVEDKRTMGXQLDRASSSLVAFSTCVIKVPLGRCEHKGKCKTC:ASRDPYCGWKEGG 540
541 ACSHLSNSRLTFFEQD:ERKNTDGLGCHNSFVALNGHSSSLLPSTTTSSTAGEV 600
541 ACSHLSNSRLTFFEQD:ERKNTDGLGCHNSFVALNGHSSSLLPSTTTSSTAGEV 600
601 SGMLDNKHLSDPSTGLGAVSSHNDKKGVTRESYKLGCHDQVPTVLLAIVLAFV 660
601 SGMLDNKHLSDPSTGLGAVSSHNDKKGVTRESYKLGCHDQVPTVLLAIVLAFV 660
661 XGAVFSGITVYCVCHHARKDVAVVQVKEKETHSRRGSGSVSVTKLSGLFGDTQSK 720
```

## RESULT 2

S-09-957-187-4

Sequence 4, Application JS/0957187

Publication No. JS200300545141

GENERAL INFORMATION:

```
APPLICANT: Shinkels, Richard A.
APPLICANT: Laroche, William
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
FILE REFERENCE: 15966-540 CIP
CURRENT APPLICATION NUMBER: US/09/957,187
CURRENT FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 65/123,667
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: 03/540,781
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/234,082
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: 60/233,798
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 60/174,485
PRIOR FILING DATE: 2000-01-04
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 939
TYPE: PRT
ORGANISM: Homo sapiens
US-09-957-187-4
```

Query Match 90.34; Score 4921; DB 11; Length 939;

Best Local Similarity 95.61; Pred. No. C;

Matches 929; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MRSALLLYPTLLHPAGAGPPEDSEP:SIHGNYTKQYVFGVGHKPGRNTTORHRLDIOM 60
DB 1 MRSALLLYPTLLHPAGAGPPEDSEP:SIHGNYTKQYVFGVGHKPGRNTTORHRLDIOM 60
QY 61 LYMNGSLYIARPHITVYD:ITSHTEETVCSKLTWKSACADVDTCRMKGKGCREF 120
DB 61 LYMNGSLYIARPHITVYD:ITSHTEETVCSKLTWKSACADVDTCRMKGKGCREF 120
QY 121 LKVLKKXDDALFYCGTNAFPCENFKNMOTLEFFDFSSQVACQYDAKHANVALFAG 180
DB 121 LKVLKKXDDALFYCGTNAFPCENFKNMOTLEFFDFSSQVACQYDAKHANVALFAG 180
QY 181 KLYSATVTDLAIDAVIYRS:GESPTLRTVAKHDKWKEKPYFQAVDGYIYFFREIA 240
DB 181 KLYSATVTDLAIDAVIYRS:GESPTLRTVAKHDKWKEKPYFQAVDGYIYFFREIA 240
QY 241 VEYTMGVVPPVAVQVCKDMGSOQVLEKQNTSF:KAPLNGSVPGDSHFVYNIQAVT 300
DB 241 VEYTMGVVPPVAVQVCKDMGSOQVLEKQNTSF:KAPLNGSVPGDSHFVYNIQAVT 300
QY 301 DVIRINGRDVVLAFTSTPYRS:IPCSAVCAVDMLDIASVFTGRPEQKSPSTWTFV 360
DB 301 DVIRINGRDVVLAFTSTPYRS:IPCSAVCAVDMLDIASVFTGRPEQKSPSTWTFV 360
QY 361 VPKRPPCCAGSSSLERYATSNFDDTLAFIKTHPLMDEAVPS:FNRPFLSTWYAYL 420
DB 361 VPKRPPCCAGSSSLERYATSNFDDTLAFIKTHPLMDEAVPS:FNRPFLSTWYAYL 420
QY 421 TKAVDTAAGPYQNTVVFILGSEKGIILKPLARIGNSGF:NDSLFLEKSVYNSEKSYD 480
DB 421 TKAVDTAAGPYQNTVVFILGSEKGIILKPLARIGNSGF:NDSLFLEKSVYNSEKSYD 480
QY 481 GVEDKRTMGXQLDRASSSLVAFSTCVIKVPLGRCEHKGKCKTC:ASRDPYCGWKEGG 540
DB 481 GVEDKRTMGXQLDRASSSLVAFSTCVIKVPLGRCEHKGKCKTC:ASRDPYCGWKEGG 540
QY 541 ACSHLSNSRLTFFEQD:ERKNTDGLGCHNSFVALNGHSSSLLPSTTTSSTAGEV 600
DB 541 ACSHLSNSRLTFFEQD:ERKNTDGLGCHNSFVALNGHSSSLLPSTTTSSTAGEV 600
QY 601 SGMLDNKHLSDPSTGLGAVSSHNDKKGVTRESYKLGCHDQVPTVLLAIVLAFV 660
DB 601 SGMLDNKHLSDPSTGLGAVSSHNDKKGVTRESYKLGCHDQVPTVLLAIVLAFV 660
QY 661 XGAVFSGITVYCVCHHARKDVAVVQVKEKETHSRRGSGSVSVTKLSGLFGDTQSK 720
```

```

661 MGAVFSGITVYVCDHRRKQVAVYQKKEKELTHSRGSMSSVTKLSGLFGDTQSKDKPPE 720
721 AILTPVXHNKGLATPGNTAKMLKADQHLDLTALPTPESTPTLQKQKPSRGSREWERK 780
722 AILTPMHNGKGLATPGNTAKMLKADQHLDLTALPTPESTPTLQKQKPSRGSREWERK 780
781 QNLINACTKMPMGSPVITDPLASASHIPSVVLP:TOOGYCHEYVDQPKXSEVAQ 840
781 QNLINACTKMPMGSPVITDPLASASHIPSVVLP:TOOGYCHEYVDQPKXSEVAQ 840
841 MALEDQAATLEYKTIKERLSSKSPNHGVLNVDLSLPKVPQREASJGPPGASLSQTGL 900
841 MALEDQAATLEYKTIKERLSSKSPNHGVLNVDLSLPKVPQREASJGPPGASLSQTGL 900
901 SKLEMHSSSYGVYDKRSYPTNSLTRSHQT 932
901 SKLEMHSSSYGVYDKRSYPTNSLTRSHLT 932

RESULT 3
S-09-957-187-6
Sequence 6, Application US/09957187
Publication No. US20030054514A1
GENERAL INFORMATION:
APPLICANT: Shinkets, Richard A.
APPLICANT: Laroche, William
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
FILE REFERENCE: 15966-540 CIP
CURRENT APPLICATION NUMBER: US/09/957,187
CURRENT FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 60/123,657
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: 09/520,731
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/234,632
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: 60/233,798
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 60/274,435
PRIOR FILING DATE: 2000-01-04
NUMBER OF SEQ ID NOS: 85
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 6
LENGTH: 884
TYPE: PRT
ORGANISM: Homo sapiens
S-09-957-187-6

Query Match 84.2%, Score 4592.5, DB 11, Length 884;
Best Local Similarity 93.7%, Pred. No. 0;
Matches 873; Conservative 2; Mismatches 2; Indels 55; Gaps 1;

1 MRSEALLVFTLHFAGAGPEDEPISIGHGNTKQYFVVGKPKGRNTTOPRR:CIQY 60
1 MRSEALLVFTLHFAGAGPEDEPISIGHGNTKQYFVVGKPKGRNTTOPRR:CIQY 60
61 IMENGLYLAARPHITVVDITHTETIYCSKLTWKSPQANDVTCRMKGKIDCHNP 120
61 IMENGLYLAARPHITVVDITHTETIYCSKLTWKSPQANDVTCRMKGKIDCHNP 120
121 IKVLLKNDALFVCGGNAPSR:RNKYNMTLEPPGDEPSSKARCPYDAKXAVVA:PAQ 180
121 IKVLLKNDALFVCGGNAPSR:RNKYNMTLEPPGDEPSSKARCPYDAKXAVVA:PAQ 180
181 KLYSATVTDLAIDAVIYRSLGE:PTLRTYKHSKWLKEPYFVQANDVGY:YFFFRRIA 240
181 KLYSATVTDLAIDAVIYRSLGE:PTLRTYKHSKWLKEPYFVQANDVGY:YFFFRRIA 240
241 VEYNTMKVVFPRVAVQCKNDMG:ISQAVLKKQNTSFLKRLNCSYFGDSHFYNILCAVT 300
241 VEYNTMKVVFPRVAVQCKNDMG:ISQAVLKKQNTSFLKRLNCSYFGDSHFYNILCAVT 300

```

```

301 DVIRINGRVYVLTAFSTPYNSIPGSAYCAYMDLIDASVFTGRFKEOKSPDSTWTFVDER 360
301 DVIRINGRVYVLTAFSTPYNSIPGSAYCAYMDLIDASVFTGRFKEOKSPDSTWTFVDER 360
361 VFKPRPGCCAGSSSSJERYATSNPEPDDTLNFIKTHPLMDCAVESINRKPWFELTMMVYRL 420
361 VFKPRPGCCAGSSSSJERYATSNPEPDDTLNFIKTHPLMDCAVESINRKPWFELTMMVYRL 420
421 TKIAYDTAAGPYQNH:TVVFLGSSKGIILKFLAIGNSGFLNDSLFLEMSVYNSEKSYD 480
421 TKIAYDTAAGPYQNH:TVVFLGSSKGIILKFLAIGNSGFLNDSLFLEMSVYNSEKSYD 480
481 GVEDKRMGMGLDSASSSLYVAFSTCVIKVPLGRCHRGKCKTCIASRFGYQGWKEGG 540
481 GVEDKRMGMGLDSASSSLYVAFSTCVIKVPLGRCHRGKCKTCIASRFGYQGWKEGG 540
541 ACSHLSFNSRATTFEOD:ERKNTDGLGCHNSFVALNGHSSSLFSTTTSESTAOEGYER 600
541 ACSHLSFNSRATTFEOD:ERKNTDGLGCHNSFVALN----- 576
601 GGMLDKWKLDELSPDSTDLGSAVSSRHQDKGVTRESYLKCHDCLVPTVLLAIAVILAFV 660
601 GGMLDKWKLDELSPDSTDLGSAVSSRHQDKGVTRESYLKCHDCLVPTVLLAIAVILAFV 660
661 MGAVFSGITVYVCDHRRKQVAVYQKKEKELTHSRGSMSSVTKLSGLFGDTQSKDKPPE 720
661 MGAVFSGITVYVCDHRRKQVAVYQKKEKELTHSRGSMSSVTKLSGLFGDTQSKDKPPE 720
721 AILTPMHNGKGLATPGNTAKMLKADQHLDLTALPTPESTPTLQKQKPSRGSREWERK 780
721 AILTPMHNGKGLATPGNTAKMLKADQHLDLTALPTPESTPTLQKQKPSRGSREWERK 780
781 QNLINACTKMPMGSPVITDPLASASHIPSVVLP:TOOGYCHEYVDQPKXSEVAQ 840
781 QNLINACTKMPMGSPVITDPLASASHIPSVVLP:TOOGYCHEYVDQPKXSEVAQ 840
841 MALEDQAATLEYKTIKERLSSKSPNHGVLNVDLSLPKVPQREASJGPPGASLSQTGL 900
841 MALEDQAATLEYKTIKERLSSKSPNHGVLNVDLSLPKVPQREASJGPPGASLSQTGL 900
901 SKLEMHSSSYGVYDKRSYPTNSLTRSHQT 932
901 SKLEMHSSSYGVYDKRSYPTNSLTRSHLT 932

RESULT 4
US-09-957-187-30
Sequence 30, Application US/09957187
Publication No. US20030054514A1
GENERAL INFORMATION:
APPLICANT: Shinkets, Richard A.
APPLICANT: Laroche, William
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
FILE REFERENCE: 15966-540 CIP
CURRENT APPLICATION NUMBER: US/09/957,187
CURRENT FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 60/123,657
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: 09/520,781
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/234,062
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: 60/233,798
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 60/174,485
NUMBER OF SEQ ID NOS: 85
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 30
LENGTH: 630
TYPE: PRT
ORGANISM: Homo sapiens
US-09-957-187-30

```

```

Query Match      61.7%   Score 3362; DB 11; Length 626;
Best Local Similarity 100.0%; Pred. No. 5,2e-274;
Matches 628; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

17 GAGFEDSEPISSHGNITKQYVFGVGHKPPRNTTQRRLDIOQIMMNGTLYIAARHIIT 78
25 1 GSGFEDSEPISSHGNITKQYVFGVGHKPPRNTTQRRLDIOQIMMNGTLYIAARHIIT 60
25 77 YVVDTSHTSEIYCSKKTWKSRQADVDTCKMKGKHQKOECHNFVKLLKNDALFVCGTN 136
25 61 YVVDTSHTSEIYCSKKTWKSRQADVDTCKMKGKHQKOECHNFVKLLKNDALFVCGTN 120
25 137 TNAPSCRNKYKMDTLEPFGDFSGMARCPYDAKHANVALFADGKLVSATVTFDAIDAV 196
25 121 TNAPSCRNKYKMDTLEPFGDFSGMARCPYDAKHANVALFADGKLVSATVTFDAIDAV 180
25 197 YVRSGESPTLRVTKHDSKMLKEPYFVQADVDTCKMKGKHQKOECHNFVKLLKNDALFVCGTN 256
25 181 YVRSGESPTLRVTKHDSKMLKEPYFVQADVDTCKMKGKHQKOECHNFVKLLKNDALFVCGTN 240
25 257 VCKNDMGGSQVLEKQWTSFLKARLNCSPGDSHFYFNILQAVTDVIRINGRDVVLATFS 326
25 241 VCKNDMGGSQVLEKQWTSFLKARLNCSPGDSHFYFNILQAVTDVIRINGRDVVLATFS 300
25 317 TPVNS:PGSAVCAYDMLDCLASVFTGRFEKQKSPDSTWTPDPRVPRPCCAGSSLE 376
25 301 TPVNS:PGSAVCAYDMLDCLASVFTGRFEKQKSPDSTWTPDPRVPRPCCAGSSLE 360
25 377 RVATSEFPDQTLNFKTHPLMDEAVPS:FNRPWFLRTMVRVRLTKIAVDTAAGPYQNTTV 436
25 361 RVATSEFPDQTLNFKTHPLMDEAVPS:FNRPWFLRTMVRVRLTKIAVDTAAGPYQNTTV 420
25 437 VVFLGSEKGIILKFLARIGNSGFLNDSLEESVYNSEKSYGVGDEKRIKNGXQJDRAS 496
25 421 VVFLGSEKGIILKFLARIGNSGFLNDSLEESVYNSEKSYGVGDEKRIKNGXQJDRAS 480
25 497 SSUYVAFSTCVIKVPLGRCERHGHKCKT:CIASRDPCYGMWKEGACSHLSNRLTFEODIE 556
25 481 SSUYVAFSTCVIKVPLGRCERHGHKCKT:CIASRDPCYGMWKEGACSHLSNRLTFEODIE 540
25 557 IBERGNTDGLGCHNSFVALNGHSSLSLPTSTTSCSTAQEGYESRGXLDKHLJLSDPST 616
25 541 IBERGNTDGLGCHNSFVALNGHSSLSLPTSTTSCSTAQEGYESRGXLDKHLJLSDPST 600
25 617 DGLAVSSRNHCKKGVIRRESYKLGHCQ 645
25 601 DGLAVSSRNHCKKGVIRRESYKLGHCQ 626

RESULT 5
US-09-957-187-63
Sequence 83, Application US/09957187
Publication No. US20030354514A
GENERAL INFORMATION:
APPLICANT: Shimkots, Richard A.
APPLICANT: LaRoche, William
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY
FILE REFERENCE: 5946-540 CIP
CURRENT APPLICATION NUMBER: US/09/957,187
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 60/232,657
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: 09/520,781
PRIOR FILING DATE: 2000-09-23
PRIOR APPLICATION NUMBER: 60/234,092
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: 60/233,795
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 60/174,435
PRIOR FILING DATE: 2000-01-04
NUMBER OF SEQ ID NOS: 95
SOFTWARE: Patent In Ver. 2.1

```

Query Match:		36.9%	Score 2043.5	DB 12	Length 888
Best Local Similarity		45.3%	Pred. No. 3 4e-163		
Matches		424	Conservative 123	Mismatches 241	Gaps 20
Y	5	ALLLYTTLJHFAAGAFSESEPSISGNYTROYFVGVHKKPGNTTOR--HRLDJQXNM 62			
P	12	ALLLELLLLGGAGHLPEDPPASVAPROVILNHYVGVSGPGRLTPASGAGDQNIQRL 71			
Y	63	INXGTLIIAARDHIYVTDJTSHTBIYNSKKTWKSRADVDIEMKSKKHDECHNPK 122			
P	72	RVRNTLFIQRDLNLYVELEPTSTLRQKRCWTMSNFSDINVCVMKQKQGEGRNPK 131			
Y	123	VLLKKHDDALFVGCTNAFNSCRJYKMDTLEPPSCDEFSYARQPYKAKHVALPADGKI 192			
b	132	VLLRRESILFVGGSNAPRVCANYSCTLQPVGNISYARQPYKPKANVALFSDGVL 191			
Y	183	YSATVTDFAIDAVYRSJGESPTLTGVHDSKWLKEPYFVQAVDVGCVYFFPFBTAVE 242			
b	192	FTATVTDFAIDAVYRSJSDRPTLTGVHDSKWLKEPYFVQAVDVGCVYFFPFBTAVE 251			
Y	243	YNTMGVVPVPRVAVQVKNMGSGCYKAKQWMSPLKALNCSVPQDSHYEYFILLQNTVM 302			
P	252	PYLKKVVSVRARVCKNDVGGSPVLEKQWTSFLKAPLNCYVGRSHYFYNLCQAVTGV 311			
Y	303	IRINGRVVLTATSTFYNSIPGSAVCAVDMLDVSVETGPFPEKQKSPDSTWVPVPERVE 362			
P	312	VSLGSRPVVLATSTFNSIFGSAVCAFLTVQNAVFSRPFPEKQKSPSTWVPVEDQVPE 371			
Y	363	KPFGCCAGSSSERVATNEFDDTJNTFKTHLMDRAVPSIFENSFWLRTNRYVLEJMK 422			
b	372	RPREGCAAFGM--QYNASSALFDDINFKVTHLMDRAVPSIGHAPWILRTNRYVLEJMK 429			
Y	423	IANTAGYCNKTVVFLQSEKSLIKELNAR--LGNSEFLNDSFLFENSVMSEKQSD 480			
P	430	VAVVGAGMGNTVTVFLGSEAGVLFKFLVRPKASTSGVSLGSVLEEFETVEPRCQSPF 489			
Y	481	GVPC--KRIMGMLDRASSLYVAFSTCVIKVPJGRCERHKKKCKIASRQPYCQWIKK 538			
b	490	GGGETQORLLSLEDDAASGGLANPFCVTVVVARCQYSGKMKVCISSCPYQGNAPD 549			
Y	539	GGACSHLSPNSRSTFCQIERGNVDSJGCHNSFVALNGHSSLLPSTTTSSTAGEGYE 598			
P	550	GSCIFLSPGTRAAFEQDVGSGASTSGJGDC----- 578			
Y	599	SRGSKLDKXHLDSFESTPLGSAVSHHCKKGVIRESYLKHQDQVLPVTLLAAVILA 598			
b	579	-----TGLLRASLSETRAGLVSNLVNLTSSVA 606			
Y	659	FVMGAVFSGITV--VCYCDHRRKDVAVVORKEK--LTHSRRGNSSVTKLSGLFGDTQSK 715			
b	607	FVVGAVVSGFVGWVGLRRERELA--RKCKEALAHGAGEAVJSVSL-----GERRAQ 560			
Y	716	DP-----KPEALITFLMNGKLTGNTAK--YLIXADQHHCLTLALTPPEST 761			
b	661	GGCGGGGGGGGAGVPEALLAPLMQV-----WAKATLGGGPHQDCSGLLPPEST 713			
Y	762	PTLQCKRKP-----SRGSEVERQNKINACTKMDPEWSPVITDLP---LRASPR 811			
P	714	P-LPQKRJLTPPHPHALGPRMTH-----GHPCLPASASSLLILIAFA 757			
Y	812	PSWVYLPITQQYCYHEYVDQPNHSEVAQMALEDQATLEYKTKERLSKSPNHNWLV 971			
P	758	APQPAPCEFTDQGLYAARPKRASGDEFTLPHASPPRRVV-----SAPTCLDIFA 811			
Y	872	ENLDSLP-PKVQKREASL-----GPPGASLSQT 898			
b	812	SAACGLPRFNSPPTGSLRPLGPHAPPAATLRRT 846			

ESULT 7  
S-09-931-836-35  
Sequence 35, Application US/09931836  
Publication No. US20030027249A1

GENERAL INFORMATION:  
 APPLICANT: Desnoyers, LLC  
 APPLICANT: Eaton, Dan L.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul G.  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Pan, James  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Wood, William F.  
 APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 FILE REFERENCE: P1330R1C1  
 CURRENT APPLICATION NUMBER: US/09/931,836  
 PRIOR FILING DATE: 2001-09-16  
 PRIOR APPLICATION NUMBER: 60/085579  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/112514  
 PRIOR FILING DATE: 1998-12-15  
 PRIOR APPLICATION NUMBER: 60/113300  
 PRIOR FILING DATE: 1998-12-22  
 PRIOR APPLICATION NUMBER: 60/113430  
 PRIOR FILING DATE: 1998-12-23  
 PRIOR APPLICATION NUMBER: 60/113605  
 PRIOR FILING DATE: 1998-12-23  
 PRIOR APPLICATION NUMBER: 60/115552  
 PRIOR FILING DATE: 1999-01-12  
 PRIOR APPLICATION NUMBER: 60/116643  
 PRIOR FILING DATE: 1999-01-22  
 PRIOR APPLICATION NUMBER: 60/125774  
 PRIOR FILING DATE: 1999-03-23  
 PRIOR APPLICATION NUMBER: 60/125778  
 PRIOR FILING DATE: 1999-03-23  
 PRIOR APPLICATION NUMBER: 60/125926  
 PRIOR FILING DATE: 1999-03-24  
 PRIOR APPLICATION NUMBER: 60/127035  
 PRIOR FILING DATE: 1999-03-31  
 PRIOR APPLICATION NUMBER: 60/127206  
 PRIOR FILING DATE: 1999-04-05  
 PRIOR APPLICATION NUMBER: 60/129122  
 PRIOR FILING DATE: 1999-04-13  
 PRIOR APPLICATION NUMBER: 60/130359  
 PRIOR FILING DATE: 1999-04-21  
 PRIOR APPLICATION NUMBER: 60/131270  
 PRIOR FILING DATE: 1999-04-27  
 PRIOR APPLICATION NUMBER: 60/131291  
 PRIOR FILING DATE: 1999-04-27  
 PRIOR APPLICATION NUMBER: 60/132371  
 PRIOR FILING DATE: 1999-05-04  
 PRIOR APPLICATION NUMBER: 60/132379  
 PRIOR FILING DATE: 1999-05-04  
 PRIOR APPLICATION NUMBER: 60/132383  
 PRIOR FILING DATE: 1999-05-04  
 PRIOR APPLICATION NUMBER: 60/135750  
 PRIOR FILING DATE: 1999-05-25  
 PRIOR APPLICATION NUMBER: 60/138166  
 PRIOR FILING DATE: 1999-06-08  
 PRIOR APPLICATION NUMBER: 60/144791  
 PRIOR FILING DATE: 1999-07-20  
 PRIOR APPLICATION NUMBER: 60/146970  
 PRIOR FILING DATE: 1999-08-03  
 PRIOR APPLICATION NUMBER: 60/162506  
 PRIOR FILING DATE: 1999-10-29  
 PRIOR APPLICATION NUMBER: 09/311832  
 PRIOR FILING DATE: 1999-05-14  
 PRIOR APPLICATION NUMBER: 09/380142

```
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/644243
PRIOR FILING DATE: 2000-04-22
PRIOR APPLICATION NUMBER: 09/747259
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/816741
PRIOR FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/854203
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/854290
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/874503
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 09/889599
PRIOR FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: 09/928,827
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US99/10733
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/US99/23851
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/50720
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: PCT/US00/05601
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US00/34956
PRIOR FILING DATE: 2000-12-23
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-23
PRIOR APPLICATION NUMBER: PCT/US01/21066
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 80
SEQ ID NO 35
LENGTH: 886
TYPE: PR
ORGANISM: Homo Sapien
IS-09-931-936-35
Query Match 36.93; Score 2009.5; DB 11; Length 886;
Best Local Similarity 45.23; Pos: 7,4e-160;
Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 207
AY 5 ALLVFTLLHFAAGAFEDSEPISSHGNYTKQYVFGVGHKPGENTGCR--HSLDQIMV 62
b 12 ALLLLLLGGGNHLLFFEPPLVAVRDYNNYFVVGSGOPULTRAGADLNLDRV 71
AY 63 IMNGLYLAARDHRYVDIDTSHTSELYCSKLLTWKSPQADNDTCRKGKKBCHNFIK 122
b 72 RVNRTLPIDSDNLYRVELESEPTSTELRYORKLWRSNPEDINVCYKKGQGEGRNFV 131
AY 123 VLLKXNDALFVCSNAPNRCRYKVKDTLEPPGDSGMARCPYDAKHANVALPADGKL 132
b 132 VLLRDESTLEFVCSNAPNRCYKVKDTLEPPGDSGMARCPYDAKHANVALPADGKL 131
AY 133 VSAVTDFLAIDAVYRSLGESPTLRTVGHDSKWLKEPYFVCAVYGDYIYFFPREIAE 242
```

PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/112514  
PRIOR FILING DATE: 1998-12-15  
PRIOR APPLICATION NUMBER: 60/113300  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: 60/113430  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/113605  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/113624  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/114140  
PRIOR FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: 60/115558  
PRIOR FILING DATE: 1999-01-22  
PRIOR APPLICATION NUMBER: 60/116043  
PRIOR FILING DATE: 1999-01-22  
PRIOR APPLICATION NUMBER: 60/125774  
PRIOR FILING DATE: 1999-03-23  
PRIOR APPLICATION NUMBER: 60/125778  
PRIOR FILING DATE: 1999-03-23  
PRIOR APPLICATION NUMBER: 60/125936  
PRIOR FILING DATE: 1999-03-24  
PRIOR APPLICATION NUMBER: 60/127035  
PRIOR FILING DATE: 1999-03-31  
PRIOR APPLICATION NUMBER: 60/127706  
PRIOR FILING DATE: 1999-04-05  
PRIOR APPLICATION NUMBER: 60/128122  
PRIOR FILING DATE: 1999-04-13  
PRIOR APPLICATION NUMBER: 60/130359  
PRIOR FILING DATE: 1999-04-21  
PRIOR APPLICATION NUMBER: 60/131270  
PRIOR FILING DATE: 1999-04-27  
PRIOR APPLICATION NUMBER: 60/131272  
PRIOR FILING DATE: 1999-04-27  
PRIOR APPLICATION NUMBER: 60/131294  
PRIOR FILING DATE: 1999-04-27  
PRIOR APPLICATION NUMBER: 60/132371  
PRIOR FILING DATE: 1999-05-04  
PRIOR APPLICATION NUMBER: 60/132379  
PRIOR FILING DATE: 1999-05-04  
PRIOR APPLICATION NUMBER: 60/132383  
PRIOR FILING DATE: 1999-05-04  
PRIOR APPLICATION NUMBER: 60/135750  
PRIOR FILING DATE: 1999-05-25  
PRIOR APPLICATION NUMBER: 60/138166  
PRIOR FILING DATE: 1999-06-06  
PRIOR APPLICATION NUMBER: 60/144794  
PRIOR FILING DATE: 1999-07-20  
PRIOR APPLICATION NUMBER: 60/146970  
PRIOR FILING DATE: 1999-08-03  
PRIOR APPLICATION NUMBER: 60/162506  
PRIOR FILING DATE: 1999-10-29  
PRIOR APPLICATION NUMBER: 09/311832  
PRIOR FILING DATE: 1999-05-14  
PRIOR APPLICATION NUMBER: 09/386142  
PRIOR FILING DATE: 1999-08-25  
PRIOR APPLICATION NUMBER: 09/644848  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: 09/747259  
PRIOR FILING DATE: 2000-12-20  
PRIOR APPLICATION NUMBER: 09/816744  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: 09/854208  
PRIOR FILING DATE: 2001-05-16  
PRIOR APPLICATION NUMBER: 09/854280  
PRIOR FILING DATE: 2001-05-10  
PRIOR APPLICATION NUMBER: 09/874503  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: 09/865539  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: 09/908,827  
PRIOR FILING DATE: 2001-07-18

PRIOR APPLICATION NUMBER: PCT/US99/10733  
PRIOR FILING DATE: 1998-05-14  
PRIOR APPLICATION NUMBER: PCT/US99/28551  
PRIOR FILING DATE: 1999-12-32  
PRIOR APPLICATION NUMBER: PCT/US99/30720  
PRIOR FILING DATE: 1999-12-22  
PRIOR APPLICATION NUMBER: PCT/US99/05601  
PRIOR FILING DATE: 2000-03-31  
PRIOR APPLICATION NUMBER: PCT/US99/05841  
PRIOR FILING DATE: 2000-03-32  
PRIOR APPLICATION NUMBER: PCT/US99/14042  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: PCT/US99/15264  
PRIOR FILING DATE: 2000-06-32  
PRIOR APPLICATION NUMBER: PCT/US99/23522  
PRIOR FILING DATE: 2000-08-23  
PRIOR APPLICATION NUMBER: PCT/US99/23329  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: PCT/US99/32678  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: PCT/US99/34956  
PRIOR FILING DATE: 2000-12-23  
PRIOR APPLICATION NUMBER: PCT/US99/06520  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: PCT/US99/17800  
PRIOR FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: PCT/US99/19692  
PRIOR FILING DATE: 2001-06-23  
PRIOR APPLICATION NUMBER: PCT/US99/21066  
PRIOR FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: PCT/US99/21735  
PRIOR FILING DATE: 2001-07-29  
NUMBER OF SEQ ID NOS: 50  
SEQ ID NO 35  
LENGTH: 888  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-035-977-35

Query Match 36.9%; Score 2009.5; DB 12; Length 888;  
Best Local Similarity 45.2%; Pred. No. 7.4e-163;  
Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 20;

QY 5 ALLLYFTLLHEAGAGFEDSEPISSHCNVTKYQVYVGHKPGHNTQR--HRLDIQMIM 62  
DB 12 ALLGELLGGAARSFPSEEPPLSVAPROYJNHVPVPGSGPGLTPAEGADJLNIQVLL 72  
QY 63 IMNGTIVIAARDHIYTVYIDTSHTEEVCSKKJWKSRQADVDTCRMKGKHKECHNFIK 122  
DB 72 RVNRTLFIGDRDNLRYVELAEPTSTERYCHKLFRSNPSPDINVCRMKGKGESECRNFVK 132  
QY 123 VLLKNDALFVCGTNAFNEFCRNKXMTLEPPGDEFGMARCPYDAKHANVALFADGKL 182  
DB 132 VLLRDESTLFVCGSNAPNPVCANYSIDTLQPVGDNISGMARCPYDPKHANVALFSDGML 191  
QY 193 YSATVTOPLADAVIYSSLGESPTLRVKHDSKULKEPYFVCAVDYGYIYFFEREIAYE 242  
DB 132 FTATVTFDAIDAVIYSLGDRPTLRVKHDSKAFKBPYPVHAVESHSVYFFEREIAYE 252  
QY 243 YNTGKVVFPFVAQVCKNMNGSGSORVLEKQWTSFLKARLNCSPVGDSSHFFPNTCAVTDV 302  
DB 252 FNYLEKVVSVAVARVCNDCVGSFVLEKQWTSFLKARLNCSPVGDSSHFFPNTCAVTDV 311  
QY 333 IRIKGRVVLATFTSPVNSPGSAVCAYDMEDIASVTFGRFKESKSPDSTWTPVDPDRVP 362  
DB 312 VSLGGRPVWLAVFSTPSNSPGSAVCAFDLTQVAAVFEGREFREKSPESIWTPVPEQVP 371  
QY 353 KPRGCCAGSSSLERYATSNFPPDCTLNFKTHPLMEANVPSIFNRPWFLRWYRYLTK 422  
DB 372 RPRGCCAARGM--CYNASSALPDIDLNFKVTHPLMEAVPESLGHAPWILRTKXHQLTR 429  
QY 423 IAVETAAGPYCNHVVVFGSEKGIILKFLAR--IGNSFNLNDSFLBEEMSVYNSEKSYD 480



```

b 430 VANDVGGKPKGKCTVWPGSSAGTVLKEJUNPNASTSGTSGLSVFLERPEFTYFPRDQURP 489
y 481 GVBD--KRIMXQLDRASSLYVAFSTCV:KVPJGRCERHKGKFKTICASRPYQOWKE 538
b 490 GGSETGGLLSLE:DAASGGLLAIFRCVVPVVARCOOYSGCKVXICISGQPYGMAHD 549
y 539 GGACSHLSPNSRPTFFGCIERGNTDGLDCHNSFVALNHSGLLITSTTSTAQOGE 599
b 550 -GSCIFUSPOTRAAFQCVASGASTSGDC----- 679
y 599 SPCMLDKHLLSDSDSTOM:GAVSSHNOCKKAVIPESYLKCHDO:WVLLAIAYILA 659
b 579 -----TGJ:RASLSSEDRAG:VSVNLLWTSVAA 606
y 659 FVNGAVPSGTV-YVCHHRKQVAVVQRKEK--LTHSRGMSVTKLSGPGCTQSK 715
b 607 FVNGAVPSGTVGFWGVRRELA--PRKDKEA:LAHGAGEAVLSVRL---GERRAQ 660
y 716 DP-----KPEAILTE:VXNGK:LATPONTAK-M:IKADQHLLDLTALPTTST 761
b 661 GPGRGGGGGGGAGVPEALLAPLMQNG-----WAKATLQGGPHLDSGLLPTFEQT 713
y 762 PTLOQKRP-----SRGSEWERNONLINACTKXKPMGSPVPTDLP---LRASPSH 811
b 714 P-LQKRLPTPHPHPHALGPRAWDH-----GHPLPASASSSSJALLAPAR 757
y 812 PSVVVJLPTQGGYCHEVUDQPKYSEVAQVALEDOATLEKTLKEHLSKSPHGMV 871
b 758 APEQPPAFGEFTFDRGYAARFGRASHGCDFTLTHASFDHRRV-----SAPTSELDPA 811
y 872 ENLDSLP-PRKVPQREASL-----GPRGASLSOT 898
b 812 SAADGLPRPWSFPPTGSLRFLGPHAFANLRT 846

```

## RESULT 9

S-10-137-870-544

Sequence 544, Application US/10137870

Publication No. US2003015893A1

## GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: DeBoer, Laura

APPLICANT: Desnoyers, Jac

APPLICANT: Pivato, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Garritsen, Mary B.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tamas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zhen

TITLE OF INVENTION: SECRETED AND TRANSXYBERANE POLYMERIDES AND NUCLEIC

FILE REFERENCE: ACIDS ENCODING THE SAME

CURRENT APPLICATION NUMBER: P33301C155

FILE REFERENCE: P33301C155

CURRENT FILING DATE: 2002-05-23

Prior Application removed - See Pat. or File Wrapper

NUMBER OF SEQ ID NOS: 550

SEQ ID NO 544

LENGTH: 888

TYPE: PRT

ORGANISM: Homo Sapien

S-10-137-870-544

Query Match

Best Local Similarity 36.9%; Score 2009.5; DB 12; Length 888;

Matches 423; Conservative 130; Mismatches 141; Gaps 20;

## RESULT 10

US-10-146-016-544

Sequence 544, Application US/10140018

Publication No. US2003013488A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

```

APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvarcoff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C158
CURRENT APPLICATION NUMBER: US/10/140,018
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
S-10-140-018-544

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7,4e-160;
Matches 423; Conservative 130; Mismatches 242; Indels 141; Gaps 20;

Y 5 ALLAYFTLLHFAGAGPDESEPTSHGNVTKQYVFGVGHKPGRNTTQR--HRLDIQIMV 62
b 12 ALLLLLLGGAGS:FFBEPPLSVAPROY.NHVPVFGSGPGLTPAEGADDLNIQRLV 71
Y 63 INMGTYIAARDHNYTVYDIDTSHTTEEYCSKLTWKVQRQADVTCRMKGKHKDECHNPIK 122
b 72 RVNRTLFIGDRDKLYRVELEPPTSTELRYQKLTWRGNPSDINVRMKKGQEGECRNPKV 131
Y 123 VLLKKDDALVCGTNAFNPSCRNKKVDTLEPFGDESGKARCPYDAKIANVALFAQCKL 182

US-10-140-021-544
Application: US/10/140,021
Publication No. US20030138886A;
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvarcoff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C167
CURRENT APPLICATION NUMBER: US/10/140,021
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
S-10-140-021-544

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred. No. 7,4e-160;
Matches 423; Conservative 130; Mismatches 242; Indels 141; Gaps 20;

Y 5 ALLAYFTLLHFAGAGPDESEPTSHGNVTKQYVFGVGHKPGRNTTQR--HRLDIQIMV 62
b 12 ALLLLLLGGAGS:FFBEPPLSVAPROY.NHVPVFGSGPGLTPAEGADDLNIQRLV 71
Y 63 INMGTYIAARDHNYTVYDIDTSHTTEEYCSKLTWKVQRQADVTCRMKGKHKDECHNPIK 122
b 72 RVNRTLFIGDRDKLYRVELEPPTSTELRYQKLTWRGNPSDINVRMKKGQEGECRNPKV 131
Y 123 VLLKKDDALVCGTNAFNPSCRNKKVDTLEPFGDESGKARCPYDAKIANVALFAQCKL 182

```

132 VLLURDESTLVGSGRAFPVCAVYSIDTLCVGENISVAVRCPYDPMANWALPDSXHL 431  
183 YSATVTDPLADAVIRSIQSPRLTAVHDSKWKKEPFPYDCHDGDYIYFFRRIAVE 442  
192 FTATVTDPLADAVIRSIQSPRLTAVHDSKWKKEPFPYDCHDGDYIYFFRRIAVE 451  
243 YNMGKVVPRVAOVCKMGGGQVLEKOWTSFLKAKLNTSVPGSHRYENILQAVTGW 402  
252 FVYLEKVVSRVAVCHDVGSGRULERXWTFUKAFJNCSVPGSHRYENILQAVTGW 411  
303 IRNGRNVVATSTPNSIPGSAVCAVYLDIASVFTDPRKSKSDSTMTVPDPRYP 462  
312 VSLGGRVWLAVESTPSNIPGSAVCAVYLDIASVFTDPRKSKSDSTMTVPDPRYP 471  
363 KPRPGCCAGSSSLERYATSNFEDTLNFKTHLMDAIVSLEKWPFLATVVRVRLTK 422  
372 RPRPGCCAGSSSLERYATSNFEDTLNFKTHLMDAIVSLEKWPFLATVVRVRLTK 431  
423 IAVDTAAGPYONHTVWFLGSEKGTIIKFPJAP--IGNSGFLVNSLFLDEKSVNSEXCSVD 480  
430 VAVDVGAGPWGNTVWFLGSEAGTVLKFVPRNASTSOTSGLSVFLBEETVPRDCCRP 489  
481 GVBD--KAINMGOLDRASSSLRYAVFTCVIKVPLGRCERHKGKKTCTIASRDPYCGWIK 538  
490 GGBETGQRLSLSLDAASGGGLAAPPFCVVRVVARCOQYSGCKNCLGSDPYCGWAPD 549  
539 GGACSHLSPNSRLTFEQDIBRGNTDGLGDCHNSFVALNGHSSSLJPTTTTSCSTAOGEYE 599  
550 -GSCIFLSPGTRAAAFECQVSGASTSGJGDC----- 578  
599 SRGGMLDKWHLLDSEFSTDPGLAVSGHNDKXGVIRESVYKGHDCQVPTVLLAIAVILA 658  
579 -----TGLRASLSEDRAGLVSVNLLVTSVAA 606  
659 FVMGAVFSGITV-YVCVDRHRKDVAVVQRKEKE--JTHSRRGSSSVTKLSGLFGDQOSK 715  
607 FVMGAVVSGFVGVNFGVJRREREJA--RRKDEKAILAHGAGEAVSVSRJ-----GERRAQ 660  
716 DP-----KPERILTPLMHNGKLAIPONTAK-XLIKADQHHLDTALPTPEST 761  
661 GPGRGSGGGGAGVPRPEALIPLMQNG-----WAKATLLOGGPHDLDOSGLLPTPEQT 713

RESULT 12

US-10-140-274-544

Sequence 544, Application US-10-140-274

Publication No. US20030143674A1

GENERAL INFORMATION:

Applicant: Baker, Kevin P.

Applicant: Beresini, Maureen

Applicant: DeForge, Laura

Applicant: Desnoyers, Luc

Applicant: Filvaroff, Ellen

Applicant: Gao, Wei-Qiang

Applicant: Gerritsen, Mary E.

Applicant: Goddard, Audrey

Applicant: Godowski, Paul J.

Applicant: Gurney, Austin J.

Applicant: Sherwood, Steven

Applicant: Smith, Victoria

Applicant: Stewart, Timothy A.

1 APPLICANT: Tomas Daniel  
2 APPLICANT: Kananabe, Colin K  
3 APPLICANT: Wood, William  
4 APPLICANT: Zhang, Zelin  
5 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
6 FILE REFERENCE: P1330R1C161  
7 CURRENT APPLICATION NUMBER: US/10/140,274  
8 PRIOR APPLICATION DATE: 2002-05-06  
9 PRIOR APPLICATION REMOVED - See File Wrapper or Pa'm  
10 SEQ ID NO 544  
11 LENGTH: 838  
12 TYPE: PRT  
13 ORGANISM: Homo Sapien  
14 US-10-140-274-544

Query Match 36.9%, Score 2009.5, DB 12, Length 838,  
Best Local Similarity 45.24, Pred. No. 7.4e-162,  
Matches 423, Conservative 130, Mismatches 241, Indels 141, Gaps 20;

QY 5 ALLVFTLLHFAGAGFPEDSEPSISHQNYTKQYVYVGHKPGRTTCR--HRJCIQMIM 62  
DB 12 ALLLLLLLGAGAGLFPBEPPLSVAFRDYLNHYVPVFGSGPGLTPAEGADDLNIQVRL 71  
QY 63 IYNGTLVIAARCHYITVDIDTSHTEESYCSKLTWKSQADVDTCKYKSKKDECHNPIK 122  
DB 72 RVNRTLFIGDRNDYRVELEPTSTELRYQRKLWRSNPSPDINCYRMKGQEGECRNPFV 131  
QY 123 VLLKNDALVCGGTNAFPNCPNRYKVDLTLEPPFDEBSGMARCPYCAKHANVALFADGKL 182  
DB 132 VLLURDESTLVGSGRAFPVCAVYSIDTLCVGENISVAVRCPYDPMANWALPDSXHL 191  
QY 183 YSATVTDPLADAVIRSIQSPRLTAVHDSKWKKEPFPYDCHDGDYIYFFRRIAVE 242  
DB 192 FTATVTDPLADAVIRSIQSPRLTAVHDSKWKKEPFPYDCHDGDYIYFFRRIAVE 251  
QY 243 YNMGKVVPRVAOVCKMGGGQVLEKOWTSFLKAKLNTSVPGSHRYENILQAVTGW 302  
DB 252 FVYLEKVVSRVAVCHDVGSGRULERXWTFUKAFJNCSVPGSHRYENILQAVTGW 311  
QY 303 IRNGRNVVATSTPNSIPGSAVCAVYLDIASVFTDPRKSKSDSTMTVPDPRYP 362  
DB 312 VSLGGRVWLAVESTPSNIPGSAVCAVYLDIASVFTDPRKSKSDSTMTVPDPRYP 371  
QY 363 KPRPGCCAGSSSLERYATSNFEDTLNFKTHLMDAIVSLEKWPFLATVVRVRLTK 422  
DB 372 RPRPGCCAGSSSLERYATSNFEDTLNFKTHLMDAIVSLEKWPFLATVVRVRLTK 429  
QY 423 IAVDTAAGPYONHTVWFLGSEKGTIIKFPJAP--IGNSGFLVNSLFLDEKSVNSEXCSVD 480  
DB 430 VAVDVGAGPWGNTVWFLGSEAGTVLKFVPRNASTSOTSGLSVFLBEETVPRDCCRP 489  
QY 481 GVBD--KAINMGOLDRASSSLRYAVFTCVIKVPLGRCERHKGKKTCTIASRDPYCGWIK 538  
DB 490 GGBETGQRLSLSLDAASGGGLAAPPFCVVRVVARCOQYSGCKNCLGSDPYCGWAPD 549  
QY 539 GGACSHLSPNSRLTFEQDIBRGNTDGLGDCHNSFVALNGHSSSLJPTTTTSCSTAOGEYE 599  
DB 550 -GSCIFLSPGTRAAAFECQVSGASTSGJGDC----- 578  
QY 599 SRGGMLDKWHLLDSEFSTDPGLAVSGHNDKXGVIRESVYKGHDCQVPTVLLAIAVILA 658  
DB 579 -----TGLRASLSEDRAGLVSVNLLVTSVAA 606  
QY 659 FVMGAVFSGITV-YVCVDRHRKDVAVVQRKEKE--JTHSRRGSSSVTKLSGLFGDQOSK 715  
DB 607 FVMGAVVSGFVGVNFGVJRREREJA--RRKDEKAILAHGAGEAVSVSRJ-----GERRAQ 660  
QY 716 DP-----KPERILTPLMHNGKLAIPONTAK-XLIKADQHHLDTALPTPEST 761  
DB 661 GPGRGSGGGGAGVPRPEALIPLMQNG-----WAKATLLOGGPHDLDOSGLLPTPEQT 713

```

762 PTJQCKRKP-----SRGSRERNKCNLINACTKMP2EMGSPVPTDLP-----LAASRSE 811
714 P-LPCKRKPTEPHPHALGERAWH-----GHLLPASASSSHLLAR 757
812 PSVVVLRITQGYOYQYEVQCPKNSVAQMALEDOATLEYKTIKHLSSKSPKNGVNLV 871
758 APEQPPAPGEPFDGRLYAARGFASHGSPPLTPHASPERRVW-----SAPTGLDPA 811
872 ENLDSLP-PKVQREASJ-----GPFQASLSCT 998
812 SAADGLPREWSPPTGSLRRPLGHPAPPAATLRRT 846

RESULT 13
S-10-140-471-544
Sequence 544, Application US/10/140.87
Publication No. US20030134354A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P33301C163
CURRENT APPLICATION NUMBER: US/10/140.471
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See File Wrapper of Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 544
LENGTH: 888
TYPE: PRT
ORGANISM: Homo Sapien
S-10-140-471-544

```

```

Query Match 36.9%; Score 2009.5; DB 12; Length 888;
Best Local Similarity 45.2%; Pred No. 7.4e+160;
Matches 423; Conservative 130; Mismatches 241; Indels 141; Gaps 20;

5 ALLYPTLHFAAGAPFSEPTISGHSNYTKOYPVFGHKPGRNTTOR--HRLDTQMLV 62
12 ALLLLLLGGAAGLFPBPPPLSVAPRYLNNYPVFGSGGR-LTPAEGADGLNQRVL 71
63 INMGTLVLAARDHIVTDITSHTEIYSKSLTKWSRQADYDTCRMKKHKHDECHNPK 122
72 RVNRTLFGRCNLYRVELEPTSTETRYCRKLTWASNSPDINVCRMKKQKQEGCRNFVK 131
123 VLLKKNDGLFVCGTWAFNPSGRNYKMDLTPEPGBDFSGNARCPYDAXHANVALPADGKL 182
132 VLLRDESLFVCGSNAPFVCANYSIDTLQPVGN--SGMARCPYDPKXANVALSDGML 191
183 YSATVTDPLAIDAVIKYS--GESPTLTVKHDSKWLKEPYFVCAVDYGVYVFFPREIAVE 242
192 FTATVTDPLAIDAVIYSLGRDPTLRVXHSKWKPEYFVHAVENGSHVYVFFPREIAVE 251
243 YNTMGKVYFPRVACVKNMGSGQRVLEKQWTSFLKARLNCQVSGESHYFVNLQAVITY 302
252 FNYLRVWVSVRVAQKNDVGGSRVLEKQWTSFLKARLNCQVSGESHYFVNLQAVITGV 311
303 IRENGRQVLA--FTSTPYS--PGSAVCAVDYLDIASVFTQHHKCKSPDSTNTFVPCSEAVF 362

```

```

312 VSLGCRPVVLAVSTPNSIPGSAVCAFDLTQVAAVEGFRBOKSPESVWTPVPEQVP 371
363 KPRPGCCAGSSSERYATSNPEFDDTLNFIKTHPLMDEAVPSIFNRPMFLRTMVRYLTK 422
372 RPRPGCCAAQPM--QYNASSALPDIILNFVKTHPLMDEAVPSLGHAPWILRTLMPHQLTR 429
423 IAVDTAAGPQNH--VWF--GSEKGIILKFLAR--IGNSGFLNDSLFJBEEMSVYNSEKSYD 480
430 YAVDVGAGPWNQ--VWF--GSEAGTVLKLVRPNASTSGTSLSVFJEEFYTYFRDRCRP 489
481 GVED--KRIMGMOULDRASSGLYVAFSTCV--KVPLGRCERHGCKKCTCIASRDPCYCKHIKE 538
430 GGETSORLISLEJLDAASGGLLAAPFCVVRVVARCOYSGCKMKNCIGSDPYCGWAPC 543
539 GGACSHLSPNSRJTFEQDIERGTGLGDCCHNSFVALNGHSSSLPFTTTSDSTAQSGYE 598
550 -GSCIFLSPGTRAAFEQOVSGASTSGLGDC----- 578
599 SRGGMLDKXLLDSPDSTDLGAVSSHHQDKGVIESYLVKGHDLQVFPVTLAIAVILA 658
579 -----TGLLRASLSEDRAGLVSNVLLVTSVAA 606
659 FWGAVFSGITV-YVCVDHRRKQVAVVQRXEKE--LTHSRRGSMSSVTKLSGLFGDQSK 715
607 FVGVAVVSGFSGVGFVGRERRELA--RRDKKAILAHGAGEAVLSVRL----GERRAQ 660
716 DP-----KFEAILTPXHNKGLATPONTAK-KLIKADQHHLDLTALFTTEST 761
661 GPGGFGGGGGGAGVPRPEALLAPLVQNG-----WAKATLLQGPCHLDSGLLFTPEOT 713
762 PTJQCKRKP-----SRGSRERNKCNLINACTKMP2EMGSPVPTDLP-----LRASPSH 811
714 P-SPKELP-TPRHFAALGERAWH-----CHPLLEFASASSSHLLAR 757
812 IPSVVVLRITQGYOYQYEVQCPKNSVAQMALEDOATLEYKTIKHLSSKSPKNGVNLV 871
758 APEQPPAPGEPFDGRLYAARGFASHGSPPLTPHASPERRVW-----SAPTGLDPA 811
872 ENLDSLP-PKVQREASJ-----GPFQASLSCT 998
812 SAADGLPREWSPPTGSLRRPLGHPAPPAATLRRT 846

```

```

RESULT 14
US-10-140-807-544
Sequence 544, Application US/10/40807
Publication No. US20030134354A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P33301C174
CURRENT APPLICATION NUMBER: US/10/140.807
CURRENT FILING DATE: 2002-05-07
Prior Application removed - See File Wrapper of Palm
NUMBER OF SEQ ID NOS: 550

```

[illegible]

QY 481 QVED--KRLMGO-LDRASSLYVAFSTCV:KYPLGRCEHCKKCKTQIASRDPYCGWIKZ 538  
DB 490 GGETGQRLSLELDAMSGGLJAFPPRCVVPVVARCOOYSGCKWNCIGSODPYCGRAPC 543  
QY 539 CCAASHLSPKSPLTFEQOIERONTDGLGCOCHNSFYALNGUSSSILAPSTTISDSTAQESYE 592  
DB 550 -GSCIFLSFG:PAAPFOVSGASTSGLGDC----- 572  
QY 599 SRGGMLDWKHLIDSPDSTPLGAVSSHHQKKGWIRESYKGHDOAVPYTLIAIAILA 659  
DB 579 -----TGLERASLSEDFAGVSVVLEVTSSVAA 606  
QY 659 FVWGAVPSGTV-VYVCDHRKCVAVWQKKE--LTHSRGSSSVTKLSGLEDDQSK 715  
DB 607 FVWGAVVSGFVGWTVGLRERDLA--RRKKEA:LAHGAGEAVLSVRL----GERRAQ 660  
QY 716 DP-----KPEAILTFLMNGKLTATPGNTAK-VLIMKADCHLDLTATBEST 761  
DB 661 GFGRGGGGGGAGVPEALLAPVONG-----WAKATILCGGPHDLSGLLFTTQCT 713  
QY 762 PTLQCKKXP-----SRGSRKWKHQNLIANNCTKCKPKMSPVIFTDLP---LRASFSH 811  
DB 714 P-LPCKRUPTEPHFPHALGPRANDH-----GHELFASAGSSLLULAPAP 767  
QY 812 IPSVVVLRITQGYQCHYVDQPKYSEVAQVALEDOATLEYKTIKEHLSSKSPKGNLV 871  
DB 753 APEQFAPGCEFTDGRLLYARKKRASHGDFELTPHASFQRRPVV-----SFTQPIDPA 911  
QY 872 ENLDSLP-PKVFQREASL-----GPFQASLSQT 998  
DB 812 SAADG:PRFWSPPPTGSLRRP:GPHAPTAAATLRT 946

Search completed: October 23, 2003, 17:23:47  
Job time : 90 secs

GenCore version 5.1.6  
Copyright (C) 1993 - 2003 Computer i.i.d.  
OM protein - Protein search, using sw mode:  
Run on: October 23, 2003, 17:09:07 ; Search time 19 Seconds  
(without alignments)  
2297.694 Million cell updates/sec  
Title: US-09-856-681-2  
Perfect score: 5450  
Sequence: 1 MRSEALLFTLLHFGASGF.....PKFSEFAPLSTSKXNDACT 1010  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 328717 seqs, 42310953 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/prodata/2/iaa/5A\_COM5.pep.\*
  - 2: /cgn2\_6/prodata/2/iaa/5A\_COM5.pep.\*
  - 3: /cgn2\_6/prodata/2/iaa/5A\_COM5.pep.\*
  - 4: /cgn2\_6/prodata/2/iaa/5A\_COM5.pep.\*
  - 5: /cgn2\_6/prodata/2/iaa/5A\_COM5.pep.\*
  - 6: /cgn2\_6/prodata/2/iaa/5A\_COM5.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2013.5	36.9	888	4	US-09-077-940A-4
2	1997	36.6	887	4	US-09-077-940A-2
3	1519.5	27.9	930	4	US-09-254-594-6
4	1428.5	26.2	929	4	US-09-254-594-3
5	969.5	17.8	730	1	US-08-121-713D-58
6	969.5	17.8	730	1	US-08-835-268-58
7	969.5	17.8	730	2	US-09-060-610-58
8	969.5	17.8	730	3	US-08-833-391-58
9	969.5	17.8	730	4	US-09-060-610-58
10	969.5	17.8	730	5	PCT-US94-1015A-58
11	880	16.1	650	1	US-08-121-713D-60
12	880	16.1	650	1	US-08-335-268-60
13	880	16.1	650	2	US-09-060-610-60
14	880	16.1	650	3	US-08-833-391-60
15	880	16.1	650	4	US-09-060-610-60
16	880	16.1	650	5	PCT-US94-1015A-60
17	852.5	15.6	655	4	US-08-556-422A-3
18	852.5	15.6	712	1	US-08-121-713D-54
19	852.5	15.6	712	1	US-08-835-268-54
20	852.5	15.6	712	2	US-09-060-610-54
21	852.5	15.6	712	3	US-08-833-391-54
22	852.5	15.6	712	4	US-09-060-610-54
23	852.5	15.6	712	5	PCT-US94-1015A-54
24	845	15.5	712	1	US-08-121-713D-64
25	845	15.5	712	1	US-08-835-268-64
26	845	15.5	712	2	US-09-060-610-64
27	845	15.5	712	3	US-08-833-391-64

26	845	15.5	712	1	US-09-060-610-64	Sequence 64, Appl
29	845	15.5	712	5	PCT-US94-1015A-64	Sequence 64, Appl
30	779.5	14.3	724	1	US-08-121-713D-62	Sequence 62, Appl
31	779.5	14.3	724	1	US-08-835-268-62	Sequence 62, Appl
32	779.5	14.3	724	2	US-09-060-610-62	Sequence 62, Appl
33	779.5	14.3	724	3	US-08-833-391-62	Sequence 62, Appl
34	779.5	14.3	724	4	US-09-060-610-62	Sequence 62, Appl
35	779.5	14.3	724	5	PCT-US94-1015A-62	Sequence 62, Appl
36	748	13.7	775	4	US-09-308-179B-1	Sequence 2, Appl
37	736.5	13.5	862	4	US-08-556-422A-4	Sequence 2, Appl
38	654.5	12.0	607	4	US-08-556-422A-4	Sequence 4, Appl
39	586	10.8	477	1	US-08-136-922-4	Sequence 2, Appl
40	539	9.9	425	4	US-08-556-422A-7	Sequence 7, Appl
41	462	8.5	295	4	US-08-556-422A-6	Sequence 6, Appl
42	437	8.0	634	3	US-09-041-236-2	Sequence 2, Appl
43	437	8.0	634	4	US-09-771-467D-2	Sequence 2, Appl
44	437	8.0	666	3	US-09-240-410-2	Sequence 2, Appl
45	356.5	6.5	606	3	US-09-041-236-4	Sequence 4, Appl

ALIGNMENTS

Result 1  
US-09-077-940A-4  
Sequence 4, Application US/09077940A  
Patent No. 6576441  
GENERAL INFORMATION:  
APPLICANT: KIMURA, Toru et al.  
TITLE OF INVENTION: NOVEL SEAPHORIN 2 AND GENE ENCODING THE SAME  
FILE REFERENCE: 0020-4434P  
CURRENT APPLICATION NUMBER: US/09/077,940A  
CURRENT FILING DATE: 1998-06-05  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: Patent version 3.1  
SEQ ID NO 4  
LENGTH: 889  
TYPE: PPT  
ORGANISM: Homo sapiens  
US-09-077-940A-4  
Query Match 36.9%, Score 2013.5, DB 4, Length 889;  
Best Local Similarity 45.3%, Pred. No. 1.7e-132;  
Matches 424, Conservative 129, Mismatches 241, Indels 141, Gaps 20;

QY	5	ALLLYFTLLHFGAGFEDSEPTISISGNTKQVPVGHKPGRNTTOR--HRJLDIGMIM 62
DB	12	ALLLLLLGLGAGGJFFEDPPPLSVAPRDYLNHYVFPVSGPGRLTPAEGADCLNIQRVL 71
QY	63	IMNGTLYIARDHIYTYDIDTSHTTEIYCSKKLTWKSROADVOTCRMKGHKDECHRFK 122
DB	72	RVNRTLFGIDRDLNLYRVELEPPSTELRYOKRLTWRSNPSDINVCRMKGOEGRNVK 131
QY	123	VLLKKNIDALFVCGTNAPNFCRNYKMDTLEPFDEFGSMARCPYDAKAKVALFADGKL 182
DB	132	VLLLRDSESTLFPVCGSNAPNFCANYSIDTLPQVGDNISGVARCPYDPKHANVALFSDGML 191
QY	183	YSATVDFSLAIDAVIYSLGESPTLRTVKDLSKWLKEPYFVCAVDYGCYIYFFRETAVE 242
DB	192	FTATVDFSLAIDAVIYSLGSDRPTLRTVKHDSKWKKEPYFVHAEVWGHSHYFFRETAIE 251
QY	243	YNTGKVVFEVAVGVCKNDVGGSGVLEKQWTSFLKARLNCSPVGDSPHFYENILCAVTDV 302
DB	252	PNYLEKVVSEVAVGVCKNDVGGSGPVLEKQWTSFLKARLNCSPVGDSPHFYENILCAVTVG 311
QY	303	IRINGRDVLTATFTSPYNS:PGSAVCAVDMLDZASVTFGRFKEQKSPDSTWTVPVDERVP 362
DB	312	VSLGSRPVLLAVFSTPSNS:PGSAVCAFDLTQVAAVFESGRFREQKSPESIWTVPVDEQVP 371
QY	363	KPRGCCAGSSSLERYATSNFEDDTLNFYKTHPLVDEAVFSPNRFELRTWVRYPLTK 422
DB	372	RPRGCCAAPGV--QYNASSALPDDILNFVKTHPLVDEAVFESGHAPKILRTLNRHQLTR 429

```
QY 425 IAVDTAGPQNHTWFGSEKGIILKFLAR--LGNSGFINDSLPFEKSYVNSEKSYD 480
D6 430 VAVDVAGAPGNQWFGSEAGNVLKPLVRENASGTSGLVFLERETTFETTFPCGCP 489
QY 481 GVFD--KRIIMQCLDRASSLSLYAFSTCV-KVPLGRGERHGKCKKCIASRDPYQWKE 538
D6 490 GGGTIGORILLSELEFASGGLLAAPFRCVVPVVARCOQSGKNKNCISGCDPYQWAPD 549
QY 539 GGAGSHUSPHERUTFEDEIRGNTDGLGCHNSFVALNGHSSLSJASTTSDTAGDGYE 598
D6 550 -GSCIFLSPGTAAFEODVSGASTSLGDC----- 578
QY 599 SRGCMLEKHLDSPOGTDPLGAVSSHNHOKKGVIRESYLKHQDQVPTLLAIAVLA 558
D6 579 -----TCLPASLSDRAGLVSYNCLVTS 604
QY 659 FVMAVRSGITV-YVCDHRRKOVAVYORKEKE--LTHSRRGSMSSVTKL----- 705
D6 607 FVMAVVSQSVGVFGLRERELA--RRDKKAILAHGGEAVLSVRLGERGT 662
QY 716 DP-----KPEALTPELXNGKATPONTAKMLIKADCHHLDLTALPTPEST 763
D6 661 GPGGRGGGGGAGVPPEALLPYCNG-----WAKATLQGRHIDSLGLPTPECT 715
QY 762 PTLCKEKP-----SRGSEKERNQNLNACTK-----DMPKXGSPVZPTD-- 802
D6 714 P-LPCKSLPTTHPHALGRANCH-----TCLPASLSDRAGLVSYNCLVTS 604
QY 812 PPSVVVPTTQGVYREYVDQKQSEVAQMALEDQAATLBYKTKEHLSKSPNHGVLV 871
D6 758 APECPAPAGEPTPGRLVYARFGASHGDFPLTHASDPDRVV-----SATTGPDPA 811
QY 872 ENLZSLP-FKVPQREASL-----GPGASLSQTL 856
D6 912 SAASGLPRKSPPTTGLRRLRPLGHPAPPAATLRRK 846

RESULT 2
US-09-077-940A-2
; Sequence 2, Application US/09077940A
; Patent No. 6576441
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru et al.
; TITLE OF INVENTION: NOVEL SEMAPHORIN 2 AND GENE ENCODING THE SAME
; FILE REFERENCE: 2020-44268
; CURRENT APPLICATION NUMBER: US/09/077,940A
; CURRENT FILING DATE: 1998-06-05
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 887
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-077-940A-2

Query Match 35.6%; Score 1997; DB 4; Length 887;
Best Local Similarity 42.7%; Filed No. 610-181;
Matches 433; Conservative 132; Mismatches 260; Intels 120; Gaps 24;

QY 2 PSEALLDYLTHPAGAPPELSPISIGNYTKQYFVYGHKGNHTQP--RLDLIQ 59
D6 10 PPALLLILLLRVTHGLFPDEPPLSVAPREYLSHYEVGSGPRGLTPAERADLHIO 69
QY 60 MINIMNGTLYAARDHVTVDLDSHTEEVCSKMLTKSRCAVDTCRKGHKDECHN 119
D6 70 RVLRVNRTLFGDQDNLSQVLEFSTSTELRYQRLTKWRSNPSDLYCRVKKGCEGRN 129
QY 120 FKVLKKNDCALFVCGTNAPNPGCRNKMOTJEFPGDEFSGVAPRCYDAKHANVALFAD 179
D6 130 FVKVLLRDESTLFVCGSNAPNPIKANSMDTLLGDNLSGVARCPYDPKHANVALFSD 189
QY 180 GKLYSATVDFLAIDAVIYRSLGSEPTLRTYKIDSKWKLPYFQAVDYGDYTYFFEREI 239
```

```
D6 190 QMLFTATVDFLAIDAVIYRSLGSDRPTLRTYKIDSKWKLPYFVHAVEMGSHVYFFEREI 249
QY 240 AVEYNTGKVFEPVAVGVKNDXGGSORVLEKQWTSFLKAKLNCOSVFGDSHFYNILQAV 299
D6 250 AMEPNYLEKVVVSSVAVKXNDVGGSFVLEKQWTSFLKAKLNCOSVFGDSHFYNILQAV 309
QY 300 TCVIRIENGRDVIATFTSPYKSIIPGSAVCAVDMLDIASVETGFPEKCKSPDSWTVPYDE 359
D6 310 TSVVSGDGRVILAVFTSPKNSIPGSAVCAFPKQVAAVFEGRPREKCKSPESWTVPYDE 369
QY 360 RVKPREPCCAGSSSLERYATSNFPPDCTLNFIKTHPMDIEAVPSIENRPMFJRTYVYR 419
D6 370 QVERPRECCOARGH--QYNASNALPDEILNFVKTHELMDEAVPSLGHSPWJATLRRHO 427
QY 420 LTKAIVDTAGPQNHTWFGSEKGIILKFLAR--LGNSGFINDSLPFEKSYVNSEK 477
D6 428 LTRVAVDVAGAPGNQWFGSEAGNVLKPLVRENASGTSGLVFLERETTFETTFPCGCP 487
QY 478 ---SYDGVEDKRIKMXQLDRASSLSLYAFSTCV-KVPLGRGERHGKCKKCIASRDPYCG 534
D6 488 GRSSSAGEMGQRLISLELDAASGGLLAAPFRCVVPVVARCOQSGKNKNCISGCDPYCG 547
QY 535 WIKEGGACSHLSPASRLTFEODI BRGNTDGLGCHNSFVALNGHSSLSJASTTSDTAG 594
D6 548 WAPD-GSCIF-LRPGTSATFEQVSGASTSLGDC----- 580
QY 595 EGYESRGMLDKHLLDSPOGTDPLGAVSSHNHOKKGVIRESYLKHQDQVPTLLAIA 654
D6 581 -----TCLPASLSDRAGLVSYNCLVTS 604
QY 655 VLFVMAVRSGITV-YVCDHRRKOVAVYORKEKE--LTHSRRGSMSSVTKL----- 705
D6 605 SVAPFVAVVSGSVGVFGLRERELA--RRDKKAILAHGGEAVLSVRLGERGT 662
QY 706 --SGJFGDTQSKDPKPEALITPELXNGKATPONTAKMLIKADCHHLDLTALPTPEST 763
D6 663 GTGGRGGAGGGPGGPPEALLPYCNG-----TKAALLHGGPHDLDSGLLPTPEQTP- 715
QY 764 LQCKEKP-----SRGSEKERNQNLNACTK-----DMPKXGSPVZPTD-- 802
D6 716 PPKSLPTTHPHALGRANCH-----TCLPASLSDRAGLVSYNCLVTS 604
QY 803 -----LPLRASPSHPVSVVLFITQGVYREYVDQKQSEVAQMALEDQAATLBYKT 856
D6 721 PERSLCARSCASHPGGFPPLP-----HASPDRRVVSAFTGPDSSSVG----- 816
QY 857 EHLSSKSPNHGVLVNLDSLP-FKVPQREASL-----GPGASLSQTLGLSKLEMHHS 909
D6 817 -----DCLFGMSPFATSSLRPGPHGPTAALRRK-----HT 849
QY 910 SSGYVDYKSYFTNSLTPSHOATLTKRNTNNSKSHLSRNCSPGRGD-NPPAPF 963
D6 850 PMSG-----EAPFGGHRFRHA-----PADSTHL---LPGTSEPTAPPVP 887

RESULT 3
US-09-054-594-6
; Sequence 6, Application US/09054594
; Patent No. 6566194
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru
; APPLICANT: KIKUCHI, Kaoru
; TITLE OF INVENTION: NOVEL SEMAPHORIN GENE: SEMAPHORIN Y
; FILE REFERENCE: 2020-4527P
; CURRENT APPLICATION NUMBER: US/09/254,594
; CURRENT FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 6
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```



```

; NAME/KEY: misc feature
; LOCATION: (1,1)
; OTHER INFORMATION: Tissue Type: Child Brain
; NAME/KEY: misc feature
; LOCATION: (1,1)
; OTHER INFORMATION: Identification Method: P for resulting peptide
JS-09-254-594-6

Query Match      27.0%  Score 1519.5  DB 4  Length 930;
Best Local Similarity 35.9%  Pred. No. 24e-135;
Matches 391; Conservative 135; Mismatches 156; Indels 189; Gaps 32;

QY 6 LLVFTLLHPAGAGPEDSEPLISGHGNTKQVPVGVGHKPGKNTTORHRLDIONLYIN 65
DB 13 LLLLSLSPH-TQAAFCQDPLFLISLQTSPLSWFPGLEDDAVAAEL-GLDQRFILN 70
QY 66 GTLYIAARDHIYVCIQTSHT-BEIVCSKLLTKWSQADVTCRMKGKHKECHNFYKVL 124
DB 71 RTLLVAARDHVFSFDLQAEDEEGSLVKNKLTWRSQ--CMENCAVRGKLTDECYNVIRVL 129
QY 125 LKXNDALFVCGTNAIPNCRNKNDLEPFFDEFSCWACQYDAKHANVALFADGKLYS 184
DB 129 VPMGQTHLACGNSFSPVCRSGYITSLQGEELSGQARCPFDATQSWVAIPFSGSLYS 366
QY 165 ATVTDFLAIDAVIYPSLGSPTTLRTVKHDSKWLKEPVYVQAVDYGVYIFFFREIAVEYN 244
DB 189 ATAADFQASDAVYVSLGPOPP-LSAKYDSKWLREPFPVVALEHGDHVYFFLEKSL-WR 246
QY 245 TMKVYVPRVAQYCKNDMGSGQVLSKQNTSFLKARLNGSVIGDSHFYFNILQAVTDV 302
DB 249 RLKGVQSFVARVCKRDMGSGSPALDRHWTSFLKRLNCSVFGDSTFYFDVLQALTGPV 366
QY 305 INGRGVVLAFTSPYKISGSACAYDMLDIASVTCREKQKSPDSTWTPVDERVP 362
DB 309 JHRSALFGVFTTCTISIPGSAVCAYFDLDEKGFEGKFKGSGSLDGAWTPVSEDRVP 366
QY 365 RPCCAGGSSLEPYATSNFFPDQTLNFITHPIMQCAVPSITNRPFLATVYVETLMA 424
DB 369 RPQSCAGVGGAALFSSRLDPLQVLFKAPHLDPVAVPTVHQ-LLTSPRALTQVA 427
QY 425 VDTAAGPYONHTVYFLGSEKGIILKFLARIGNSGFUNDLSFLSEMSVYNSEKSVGGVGD 484
DB 428 VDMGAGPHSNITVYFLGSGDGTVLKVTGGRSGG-PEPILAEIDAVSPARCSGKRTAQ 486
QY 485 --KRIKXVCLDRASSLYVAESTCVIKVPLGRCEKCKKCTASRDYCYCKWKEGSGC 542
DB 487 TARKITGLELTGTGHLFVAFSGCIWPLSLRCARHAGACORSCASQCPYCKHSGR-C 545
QY 543 SHLSPKSLRTFEQDIERGNTDGL--GDCHNSFVALNGHSSLLPSTTSDTAQEGYESR 600
DB 546 VDIRSGSGTDVQ--AGNCSSVEHGDCG-----ATGSGSGPGDSAY--- 586
QY 601 GGMLOWKHLDSPDSTDEPLQAVSSNHQDKGVIRSGYKGDHQLVPTLLAIANVLAPV 660
DB 587 -----GYRRDLPPASASRSVPILPILLASVAAFA 615
QY 661 MGAVFSGITVYCYCHERKDVAVVQKEXELTSPRG-----SMSSVYKLSGL 708
DB 616 LGASVGLYVSCAC--RR-----AARRPKQDIEPGVPRPJSLSLARLHG- 659
QY 709 FGDQSKDKPK--EALITFMHNGKILATGNTAKMLIKADQHHLELTAPTPESTFTLQ 766
DB 660 -GGPEPPPPSKDQDAVCTPOLYTFEPPPEGVPPP-----ELACLPTPESTPELPV 709
QY 767 KRKPSSGSRSEWERNQNLINACTKMPMPGSPVPTDLP-RAQPSH-----IPSVVLPIT 821
DB 710 KHLRAAGD-PWEWKORRNA-----KEGPGPSRSGHAGGAPRVLVPP-- 752
QY 822 QCGYQHEVYDCPXMSEVAOMLEQDQATLEYKTKKEL-----SSKSPNHGYNVENLDS- 876
DB 753 -----PPPGCGG-----AVEVTTLEELJYILKGFQPRKAGSPAPLFSR 793
QY 877 -JPPKVPQREASLGPFGASLSCGLSKRLEMHSSSYGVYKRSYVNTGSLTASHCATLK 935
DB 794 ALPPE--PAPALLGSPSPREHECAEAPLRLDV-----PDEGRCAAPA----- 833
QY 936 RNNNTSSNSHL-----SRNSQFGRGDNPPAPORVPSIQVHSSQPSQAVTVSRQPS--- 988
DB 934 -RPALSAFAPRLGVGGGRLLPFSCHRAPPALLTRV-----PSGSPSRYSGGPSKHL 883
QY 989 --LNAYNSLTPRSLKRTPSJKPVPVPPKPSF-APLSRSMKN-1026
DB 864 LYLGRPEYEGRALKRVDEKPCLSLKPPLVGPSSRQAVPN 924

RESULT 4
US-09-254-594-3
; Sequence 3, Application US/09254594
; Patent No. 6566094
; GENERAL INFORMATION:
; APPLICANT: KIMURA, Toru
; APPLICANT: KIKUCHI, Kaoru
; TITLE OF INVENTION: NOVEL SEXAPHORIN GENE: SEXAPHORIN Y
; FILE REFERENCE: 0030-45279
; CURRENT APPLICATION NUMBER: US/09/254,594
; CURRENT FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 928
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1,1)
; OTHER INFORMATION: Tissue Type: Brain
; NAME/KEY: misc feature
; LOCATION: (1,1)
; OTHER INFORMATION: Identification Method: P for resulting peptide
US-09-254-594-3

Query Match      26.2%  Score 1428.5  DB 4  Length 929;
Best Local Similarity 34.0%  Pred. No. 11e-126;
Matches 360; Conservative 147; Mismatches 355; Indels 197; Gaps 32;

QY 6 LLVFTLLHPAGAGPEDSEPLISGHGNTKQVPVGVGHKPGKNTTORHRLDIONLYIN 65
DB 13 LLLLSLSPH-TQAAFCQDPLFLISLQTSPLSWFPGLEDDAVAAEL-GLDQRFILN 69
QY 66 GTLYIAARDHIYVCIQTSHT-BEIVCSKLLTKWSQADVTCRMKGKHKECHNFYKVL 124
DB 70 RTLLVAARDHVFSFDLQAEDEEGSLVKNKLTWRSQ--CMENCAVRGKLTDECYNVIRVL 127
QY 125 LKXNDALFVCGTNAIPNCRNKNDLEPFFDEFSCWACQYDAKHANVALFADGKLYS 184
DB 128 VPMGQTHLACGNSFSPVCRSGYITSLQGEELSGQARCPFDATQSWVAIPFSGSLYS 187
QY 185 ATVTDFLAIDAVIYPSLGSPTTLRTVKHDSKWLKEPVYVQAVDYGVYIFFFREIAVEYN 244
DB 188 ATAADFQASDAVYVSLGPOPP-LSAKYDSKWLREPFPVVALEHGDHVYFFLEKSL-WR 246
QY 245 T--MGKVYVPRVAQYCKNDMGSGQVLSKQNTSFLKARLNGSVFGDSHFYFNILQAVTDV 302
DB 247 TPGUGRVQFSRVAARVCKRDMGSGSPALDRHWTSFLKRLNCSVFGDSTFYFDVLQSLTG 366
QY 303 IRINGRDVLAFTSTPNYSIPGSAVCAYDMLDIASVTFGRFKSQKSPDSTWTPVDERVP 362
DB 307 VNLGRSALFGVFTTCTISIPGSAVCAYFDLDEKGFEGKFKGSGSLDGAWTPVSEDKVP 366
QY 363 KPRPGCCAGGSSLSERVATSNFPPDQTLNF-KTHPMDAEAVPSIFNRPWFJRTVYVRLTK 422
DB 367 SPRGSCAGVGAALFSSRLDPLQVLFKAPHLDPVAVPTVHQ-LLTSPRALTQVA 425
QY 423 IAVTAAAGPYONHTVYFLGSEKGIILKFLARIGNSGFUNDLSFLSEMSVYNSEKSVGGV 482
DB 426 VAVDMAGPHSNITVYFLGSGDGTVLKVLPP-GQSLGPEPILAEIDAYSHARCS--GK 482
```

```

483 QY 483 ED-- KRIMGQJQDRASSLYVAFSTCVIKVPLGRCERKCKKTCIASRDPYQCHIE 338
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
483 Db 483 RSPRAARIIGLEETGCHLFAVAFGCIYVLSRCARHAGACORCLASLDYQCHRF 342
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
539 QY 539 GGACSH:SPASPAATFEODIE-KGNTDZL--GCGHSPVALKHSLSLSTTSTSTAGF 346
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
543 Db 543 RGVNIRGPGS-----TDVHILGCEKHEKD-QCG-----ATGQSPPGFS 344
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
596 QY 596 GYESGQXLDWKHLDSFDSSTFLGAVSSHHQKGVFEESTLKHQDQVPTLLAIAY 355
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
598 Db 598 AY-----GVSRCLSPASASRSPIFILLAY 310
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
656 QY 656 ILAFVMGAVSGITVYVCCDEHKEKAVAVORKEKE-TSPSGMSSTVKLSGLFGTQSK 315
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
611 Db 611 AAFALGASVSLVSCAC--RRAN-----RRSKLETPLRPLSLRLARLHGGGEPF 364
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
716 QY 716 DKXP---EAILTFMXNGKLAIPONTAKMLKADQHELDLTALPTPSTPTLQCKEKP 372
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
665 Db 665 PPPKDCAACTPOLYTF-LPEGGSGP-----FLACLPSTTPELPVKELRAS 315
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
773 QY 773 GSEKENQNLINACTKNDPWPSPVPTDLP---LRASGHIPSVVVFITQQ---GYC 326
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
716 Db 716 GG-PWEMWONGNASEGPGPRGSCSAAGSPAPRAVVRPPPGCGGGEVEVITLLEELRYL 374
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
927 QY 927 H-----EYVDQEKYSEVAYALSDQAATLEYKTKHLSKSKENHGYNLV 371
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
775 Db 775 HGQPCKGSEPLASAPFTSPASPGALFVD-----SSPFR----- 314
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
972 QY 972 ENLDSUPP-----KVFREASLPGPGASLSQTLGSKALEMHSSSYGVYKRY 320
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
815 Db 815 ---JCVPFLRLDVPDQKRAAPSGRPALSAPARLGVSG-SRRL-----PF 356
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
921 QY 921 PTNSLRTSHQATLTKRNTNTHSSNSSLRNQSPRGCGNPPAPQVDS---IQVHSQPSG 373
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
857 Db 857 PT-----HRA-----FPGLTRYSPGGRYSGQGR 363
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
979 QY 979 QAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDV--PPKPS 305
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
934 Db 934 HLLYLGR-PGCHGRSLKAVDVKSPISPKPLATPQPA 321
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
US-08-121-713D-58
: Sequence 58, Application US/08121713D
: Patent No. 563856
: GENERAL INFORMATION:
: APPLICANT: Goodman, Corey S.
: APPLICANT: Kolodkin, Alex L.
: APPLICANT: Matthews, David
: APPLICANT: Bentley, David R.
: APPLICANT: O'Connor, Timothy
: TITLE OF INVENTION: The Semaphore Gene Family
: NUMBER OF SEQUENCES: 100
: CORRESPONDENCE ADDRESS:
: STREET: 268 Bush Street, Suite 3200
: CITY: San Francisco
: STATE: CA
: COUNTRY: USA
: ZIP: 94134
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/121.713D
: FILING DATE: 13-SEP-1993
: CLASSIFICATION: 5.4
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman, Richard A.
: REGISTRATION NUMBER: 36,627

```

```

: REFERENCE/DOCKET NUMBER: B94-302-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 343-4342
: TELEFAX: (415) 343-4342
: TELEX:
: INFORMATION FOR SEQ ID NO: 58:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 730 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: Protein
: US-08-121-713D-58

Query Match 17.3%, Score 369.5, DB 1, Length 730;
Best Local Similarity 34.4%, Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 117; Gaps 29;

QY 5 LVFTLLHFAGAGFPEDSEFIS:SHGNVTQKYPVFVGHKPGKNTTCR-----HSLDQMI 61
DB 11 LNWALH-AAANWDVSP-----KMVQF-----GEERVQFJGNESHKDHFKLL 54

QY 62 NEMGTLLVIAAPDHIYTVQID--TSHTTEIYCSKLTWKSRQADVDTCRMKGKFKDECHN 119
DB 55 EKCHNSLLVGARNIVYANISLPDLTFTEQ-----RISWHSSGAHREL CYLKGXSDDCQN 109

QY 120 FIKVLLKKNDDALFVCGTNAFNPSCRNYKMDLTLPFGD-----SPSGMARCPYDAKANV 174
DB 110 YIRVLAKIDDDRVLCGTNAVYKPLCRHYALKD---GDYVVEKEYEGRGCLCPFPDCHNST 165

QY 175 ALFADGKLYSATVDFLAIDAVIYRSLGESPTLRTVKHSKWLKEPYFVQAVDYGDYIYF 234
DB 166 AYSEGGLYSATVADFSCDTLIIYRG-----PJTERSDLKQLNAPNVTNMEYNDPIFP 220

QY 235 FFRZAEVYNTMGVFRVAVQVCKNDYGGQSRVLEKQWTSFLKARLNCVSPGDSHPYFN 294
DB 221 FFERATAVEYVCGKATVSRVAVCVCKDGPHQGGDR-W-SFLKSRNLNCVSPGDPYFVN 279

QY 295 ILQAVTVIRN-N-GRDY---VLATESPYNSIPGSAVCAYDMLDIASVFTGREFKEQSPD 350
DB 290 EQSTSDILEGNVGQVEKLIYGVFTIPVNSIGSAVCAFSMKSLSEFDGPFKEQETYN 339

QY 351 STWTPVDPDERVPKPRGCCAGSSSDERYATSNRPPDQTLNFIKTHPLMCGAVPPIPNRPW 410
DB 340 SNWLAVPSLKVPSPRQCVND-----SRCLPQVSNVAFVKSHLMDAEAVZFFTFPI 391

QY 411 FLRTVRYRLTKIAYD---TAAGPYQNTTVFLIGSEKGILKFLARIGNSGFYN----- 461
DB 392 LIR-SLOYRFTKIAYDQVKTPDG--KAYDVLFITDGGKVIKAL-----NSASFSDDTV 445

QY 462 DSLPLEMSVYNSEKSGVGDVEDKRIYGMQD-RASSSLYVAFSTCV:KVPILGR- ERHG 519
DB 446 DSVVIBELQVLP-----PGVFWXNVYVRVGDGDKLVVYSDDELLAIKJHRGSGSKIT 499

QY 520 KCKTCTIASHRDPYCNKIKERGACSHL-SPN---SRJTFEQDIERGNTDGLGDCNHSVAL 575
DB 500 NCRE-CVSLQDPYCAWKNVELKCTAVGSPDMSAGKRRFIQNSLGEHKA-CGGRQTEIV- 557

QY 576 NCHSSSLPSTTTSDSTA-----QGVESRSGKMLDKHLLJSDPSDSTDPLGAVSSHHQ 628
DB 558 ----ASPVTQPTTKSSGDPVMS:HQAEFEPE---DNEIIVIGVDDSNVIPNTLAEINRA 610

QY 629 DKKGVIRESV:KGHDQ:VVPV---TJ-LAI-----AVILAFVNGAVFS 666
DB 611 GSK-----LPSQEKLPYTAETJTAIVTSCLGAVVGVGFSIGFLFS 652

RESULT 6
US-08-835-268-58
: Sequence 58, Application US/08835268
: Patent No. 5807826
: GENERAL INFORMATION:
: APPLICANT: Goodman, Corey S.
: APPLICANT: Kolodkin, Alex L.

```

APPLICANT: Matthes, David  
APPLICANT: Bentley, David R.  
APPLICANT: O'Connor, Timothy  
TITLE OF INVENTION: The Semaphorin Gene Family  
NUMBER OF SEQUENCES: 100  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 Bush Street, Suite 3200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.2b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/835,268  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,713  
FILING DATE: 13-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Ogan, Richard A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: B94-002-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
TELEX:  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 730 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-835-268-58

Query Match 17.8%; Score 969.5; DB 1; Length 730;  
Best Local Similarity 34.4%; Pred. No. 4.5e-83;  
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 23;  
QY 8 LVFTLLHFAAGAFPEDESEISHSNVTQYVFGVGHKPGKRNITQR-----HRLDQIM 61  
DB 11 LKVALH--AAAWNDVSP-----RMVGF-----GSEPVCFGLNESHKORFKLS 54  
QY 62 MINMGTVIAARDHIVTDID--TSHTTEIYCSKLTWKSROADVTCRMKGKHOECN 119  
DB 55 EKDNLSLVGARNIVNISRLDTEFEQ-----RIENHSSGARELCYLKSKSDDCN 109  
QY 120 FIKVLKXDDA-FVCGTNAFNSCRNYKMTLEPGD-----EFGNARCPYDAKHAHV 174  
DB 110 VIRVLAKIDDDRLICOTRAYKPLCRHVALKD---GVVVEKEVEGRGLCFDPCHNST 165  
QY 175 ALPADGKLVSATVTDLAIDAVIYRS:GSPST:RTVYHDSK:KBPYFQAVCYGY:YF 234  
DB 166 AIYSEGLYSATVADFSG:DLPIYR:-----PLRTESDLQ:LNAPNFVNTKEYNDF:FF 220  
QY 235 FFEZ:AVENTMGSVVFPRAVACVKNDVMSORVLEKOWTSP:KARLNC:SVPGDSHFFYN 294  
DB 221 FRETAVEINCKAIYSARVAVCKKGGPGHGGGR-KTSPKSLNCSVPGDYPFFN 279  
QY 295 ILCAVTDVIRIN-GRGV---V:ATFTSTPYN:IFGSVAVCA:YKLDIASVFTGRFKQKSP 350  
DB 280 E:QSTSDILEGNVGGQVEKLIYVFTIPVNS:IGGSVAVCAFSKKS:ILESFGGPFKEQETWY 339  
QY 351 STWTPDVERVAPRGCCAGSSSJERYATSNFFDCTLNF:KTP:KDEAVFSP:PNRW 410  
DB 340 SNMLAVFS:KVPSPRGQCVCND-----SRTLPOVS:VYFKSHT:KXDEAVFAFTRI 391  
QY 411 FLRTMVRRLTKIADV-----TAAGPYCNHTVFLGSEKGIILKFLAR:IGNSGFLN----- 461

DB 392 LIRISLQYRRTKIADVCOVR:PDG--KAYDVLEGTDDGKVIKAL-----NSASFSSTV 445  
QY 462 DSLP-BEMSVYNEKCSYGVVEDKRIYMGOLDRASSSLYVAFSTCVIKVPLGRG--ERHG 519  
DB 446 DSVV:EEBLQVLP-----PGVPKVLVYVWRMGDSKLVVVSDEILAKLHRCGSKIT 499  
QY 520 KCKYTCIASRDPYCGK:KEGCACSHL:SPN---SRLTFEQD:ERGNTDG:GDCHNSFVAL 575  
DB 500 KCRE-CVSLQDPYCAWNVKCTAVGSPDMSAGRRFIONISLGEHKACGGRPQTIV- 557  
QY 576 NGHSSLLPSTTSDSTA-----QEGYESRGYLDWKHLLDSDPTDPLGAVSSHNHQ 628  
DB 558 ---ASPVTQPTTKSGGDPVHSIHQAEFEPE---IDNEIVIGVCDNSV:PNLIAEINHA 610  
QY 629 EKKGVIRESV:KGHDLQVYV---TL-LAI-----AVLAFVYMGAVFS 666  
DB 611 GSK-----LPSQEKLPYVTAETLTIAVTSCGALVVGFS:SGFLFS 652  
RESULT 7  
US-09-060-692-58  
Sequence 58, Application US/09060692  
Patent No. 5935865  
GENERAL INFORMATION:  
APPLICANT: Goodman, Corey S.  
APPLICANT: Koidekin, Alex L.  
APPLICANT: Matthes, David  
APPLICANT: Bentley, David R.  
APPLICANT: O'Connor, Timothy  
TITLE OF INVENTION: The Semaphorin Gene Family  
NUMBER OF SEQUENCES: 100  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 Bush Street, Suite 3200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/360,692  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,713  
FILING DATE: 13-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Ogan, Richard A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: B94-002-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
TELEX:  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 730 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-060-692-58

Query Match 17.8%; Score 969.5; DB 2; Length 730;  
Best Local Similarity 34.4%; Pred. No. 4.5e-83;  
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;  
QY 6 LVFTLLHFAAGAFPEDESEISHSNVTQYVFGVGHKPGKRNITQR-----HRLDQIM 61

```
DB 11 LKVALH--AAANQNVSP-----KMYVQF-----QSERVQRPFGNESHKQHFLL 54
QY 62 MINNGTLVIAARSHIYTVD--TSHTTEIYCSKLTWKERQADVTQRMKGKHKDECHN 119
DB 55 EKDHNSLLVAGARIVVNISRLUTEFEQ-----RIEWHSSGAHRELQYKKGSEDDQON 103
QY 120 FIKVLLKKNDAFVQCTNAFNSCRNYKMDTLEFFQJ-----EFSGMARCPYDAKHANV 174
DB 110 YIRVLAKIDDERVLCGTVNAVKPLCRHVALKD---GVYVVEKEVEGRGLCPDPDQNST 165
QY 175 ALPADGKLYSATVDFLAIDAVIYRSLGESPTLTUVKHSKWLKEPYFVQAVDYGVYIF 234
DB 166 AIYSEGQYSATVADFSGDTPLIYRG-----PLRTSRDLKQLNAPNFVNTMEYNDPIFF 220
QY 235 FFRFAVEYNTMGVPRVAQVCKNDMGSGQRVLEKQWTSFLKARLNCVPGDQSHFYN 294
DB 221 FFRFAVEYNTMGKAIYSRVARVCKHKGPHQGGDR-WTSFLKSRJLNCVPGDQSHFYN 279
QY 295 ILQAVTDVIRIN-GRDY---VLATFTPYNSIPGSAVCAYDVXDIDIASVFTGRFKEQKSPD 350
DB 290 EIQSTSDIIEGNYGGQVEKLIYGVFTTPVNSIGSSAVCAFSMKLSILESFDGPFKEQETMN 339
QY 340 SMKLAVPESLKVPEPRQCCVND-----SRTLPOVSVNPFVKSHTLNDQAVPAFFTRPI 391
DB 411 PLRTMVAIRLTKIAVD---TAAGFYQNHVTVFSGSEKGIILKFLAPLIGNSGFJN----- 461
DB 332 LIRSLQVRFPTKAVQCVQVATPDG--KAYDVLFITGTDGKVIKAL-----NSAFSDSDTV 445
QY 462 DCLFLEMSVYNSEKSYDGVDERKRMNGXGLDRASSSLYVAFSTCVIKVFLGRG--ERHG 519
DB 446 DSVVTEELQVDF-----PGVPVKNLVYVWRMCGDSKLVVVSDEILAKLHRGSGSKIT 499
QY 520 KCKKTCIASRDPYCGWKEGAGCSHL-SPN---SRLTFEGDIERGNTDGLGDCHNSFVAL 575
DB 500 NCRE-CVSLQFPYCAWNVKCTAVGSPDWSACKGRFFIQNISLGEHKACGGRPQTEIV- 557
QY 629 DKGKVIRESYLKGRDQAVPV---TL-LAI-----AVILAFVYGVAFS 666
DB 611 GSK-----LPSSQEKLPYTAETLTIAIVTSLGALAVGVF-SGFLFS 652

RESULT 8
US-08-833-391-58
: Sequence 58, Application US/58033391
: Patent No. 631378:
: GENERAL INFORMATION:
: APPLICANT: Goodman, Corey S.
: APPLICANT: Klocklin, Alex L.
: APPLICANT: Matthews, David
: APPLICANT: Bentley, David R.
: APPLICANT: O'Connor, Timothy
: TITLE OF INVENTION: the Semaphorin Gene Family
: NUMBER OF SEQUENCES: 100
: CORRESPONDENCE ADDRESS:
: ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
: STREET: 268 Bush Street, Suite 1200
: CITY: San Francisco
: STATE: CA
: COUNTRY: USA
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/833,391
```

```
Query Match 17.8%, Score 969.5; DB 3; Length 730;
Best Local Similarity 34.4%; Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

QY 8 LYFTLLJHFACGPDSEDPIS--SHGNYTKOYYPVGHKPGORNTTQR-----HRLDIQMI 61
DB 11 LKVALH--AAANQNVSP-----KMYVQF-----QSERVQRPFGNESHKQHFLL 54
QY 62 MINNGTLVIAARSHIYTVD--TSHTTEIYCSKLTWKERQADVTQRMKGKHKDECHN 119
DB 55 EKDHNSLLVAGARIVVNISRLUTEFEQ-----RIEWHSSGAHRELQYKKGSEDDQON 109
QY 120 FIKVLLKKNDAFVQCTNAFNSCRNYKMDTLEFFQJ-----EFSGMARCPYDAKHANV 174
DB 110 YIRVLAKIDDERVLCGTVNAVKPLCRHVALKD---GVYVVEKEVEGRGLCPDPDQNST 165
QY 175 ALPADGKLYSATVDFLAIDAVIYRSLGESPTLTUVKHSKWLKEPYFVQAVDYGVYIF 234
DB 166 AIYSEGQYSATVADFSGDTPLIYRG-----PLRTSRDLKQLNAPNFVNTMEYNDPIFF 220
QY 235 FFRFAVEYNTMGVPRVAQVCKNDMGSGQRVLEKQWTSFLKARLNCVPGDQSHFYN 294
DB 221 FFRFAVEYNTMGKAIYSRVARVCKHKGPHQGGDR-WTSFLKSRJLNCVPGDQSHFYN 279
QY 295 ILQAVTDVIRIN-GRDY---VLATFTPYNSIPGSAVCAYDVXDIDIASVFTGRFKEQKSPD 350
DB 290 EIQSTSDIIEGNYGGQVEKLIYGVFTTPVNSIGSSAVCAFSMKLSILESFDGPFKEQETMN 339
QY 340 SMKLAVPESLKVPEPRQCCVND-----SRTLPOVSVNPFVKSHTLNDQAVPAFFTRPI 391
DB 411 PLRTMVAIRLTKIAVD---TAAGFYQNHVTVFSGSEKGIILKFLAPLIGNSGFJN----- 461
DB 332 LIRSLQVRFPTKAVQCVQVATPDG--KAYDVLFITGTDGKVIKAL-----NSAFSDSDTV 445
QY 462 DCLFLEMSVYNSEKSYDGVDERKRMNGXGLDRASSSLYVAFSTCVIKVFLGRG--ERHG 519
DB 446 DSVVTEELQVDF-----PGVPVKNLVYVWRMCGDSKLVVVSDEILAKLHRGSGSKIT 499
QY 520 KCKKTCIASRDPYCGWKEGAGCSHL-SPN---SRLTFEGDIERGNTDGLGDCHNSFVAL 575
DB 500 NCRE-CVSLQFPYCAWNVKCTAVGSPDWSACKGRFFIQNISLGEHKACGGRPQTEIV- 557
QY 576 NGRSSLLPSTTTSDSTA-----QEGYRSRGMLDKWLQSLDPSDTPDLAGVSSHNHQ 628
DB 558 ---ASPVTGPTTKSSGCPGHSIQAEFEPE---IDNE-VIGVDDSWIPNTLAEINHA 610
QY 629 DKGKVIRESYLKGRDQAVPV---TL-LAI-----AVILAFVYGVAFS 666
DB 611 GSK-----LPSSQEKLPYTAETLTIAIVTSLGALAVGVF-SGFLFS 652
```

```

RESULT 9
US-09-856-610-58
: Sequence 58, Application US/090606:0
: Patent No. 6344544
: GENERAL INFORMATION:
: APPLICANT: Goodman, Corey S.
: APPLICANT: Kolodkin, Alex L.
: APPLICANT: Mathes, David
: APPLICANT: Bentley, David R.
: APPLICANT: O'Connor, Timothy
: TITLE OF INVENTION: The Semaphorin Gene Family
: NUMBER OF SEQUENCES: 120
: CORRESPONDENCE ADDRESS:
: ADDRESSES: SCIENCE & TECHNOLOGY LAW GROUP
: STREET: 268 Bush Street, Suite 3200
: CITY: San Francisco
: STATE: CA
: COUNTRY: USA
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-PCS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/060,610
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/035,269
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman, Richard A.
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: B94-002-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 343-4341
: TELEFAX: (415) 343-4342
: TELEX:
: INFORMATION FOR SEQ ID NO: 58:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 730 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-856-610-58

Query Match 17.8%; Score 969.5; DB 4; Length 730;
Best Local Similarity 34.4%; Pred. No. 4.5e-83;
Matches 244; Conservative 124; Mismatches 224; Indels 117; Gaps 29;

Qy 8 LYFTLLHFAGAGFPESEP-SISHGNYTKQYVPVGHKPGRNTOR-----HRLDIQMI 61
Db 11 LLWVALH--AAAVNDVSP-----KNVQCF-----GEERVQFLGNESHKQHEKLS 54
Qy 62 MINMGLYIAARCHIVTDID--TSHTEEVYCSKLTWRSQRADVDTCKMKGKDECHN 119
Db 55 EKDNHLLVGARIVNISIRJCTTEFEQ-----RIEWHSSGAERELCYLKGSEDDQK 109
Qy 120 FIKVLKKNDDALFVCSNAPNSCRNKKYKDTLEPFGD-----EFGMARCPYDAKHANV 174
Db 110 YIRVLAKIDDRVLICGTNAYKPLCRHYALKD---GCVVWEKEYSGRLCFFDPDHNST 163
Qy 175 ALPADGKLYSATVTDPLAIDAVIYRS-GHSPTLRTVYHDSKWKKEBYFVOAVDYGVIYF 234
Db 166 AIVSEGLYSATVADESSTDEPLIVRG-----PLRTERSDLKQLNAPNFVNTWEYNDIEP 220
Qy 235 FFEINAVEYNTGKVIYPPRVACVCHNDMGSCORVLEKNCSTFIKALNSGVPGDSHFYFN 294
Db 221 FFEINAVEYNTGKVIYPPRVACVCHNDMGSCORVLEKNCSTFIKALNSGVPGDSHFYFN 279
Qy 295 ILOAVTDVIRIK-GRDV---VLATFTSPYKSIKSGAVCAVDNLDIASVFTGFEKCKSPC 350

```

50 11 LKVALH--BAAWNVDP-----KNYVDF-----GEEVQVFLGNSEKZKFXL 34  
QY 62 YIMNGTLYAARHIYVDID--TSHTERYCKKLTWMSRQAVDTCTRYKGRHKBCHN 119  
DB 55 EKOHNSLUGARNIVYINISLRLTFTFEQ-----R-EMWSSGAHRELQYKKGSEDCCN 109  
QY 120 FIKVLKKRDALFVCGTNAPNSCNYKVTJEFEG-----EFSOMARCPYAKHAY 174  
DB 110 YIRVLAKIDSRVLICGTINAYKLCRYALKO-----GVVYKEXYEGRLCPDFDHST 169  
QY 175 ALFADKLYSATVDFALDAIVYRSLGHSPTCTYVKHESKMKKEPYVCAVGYGVYIF 214  
DB 166 AIVSEQLYATVACFGSDPLLYRG-----PCTEBSLKOLNAPVYVYVNEVDPIFP 229  
QY 235 FFEIAYEYTNKVVPPRVAQYCKXMDGGSQVLEKOWTFPKAHLNCSVPCDSFEYFN 294  
DB 221 FFEIAYEYTNKKAIVSRVAPVCKDKGPHQGGH-WTSPKSPKNSVNDYDPVFN 279  
QY 295 ILOAVTDVIRN-GRV-----VLATESPPNSIGSAVCAVDMGLIASVNTGRKCKSPD 380  
DB 280 EIGSTSLICGKYGQVKKLYGVFTTVPNSIGSAVCAVPSKYLIESPDGFKKCEYHML 319  
QY 351 STWTFEDERVPFRPCCAGSSSJERYATSNRFPDCTIANFMTPLMDPRAVPSIFNRPM 410  
DB 340 SNMLAVPSLVKVPBPQCCVND-----SRLPQVSNPVKWSHTLYDSAVPAFTPI 391  
QY 411 FLRTMVRILTKIAYD-----TAAGPYQNHVTVFLGSEKGLILKFIARIGNSFLN----- 461  
DB 392 LIRISLOVRETKIAYDQOVRTPG--KAVDVAFIGTCDKXVIAL-----NSASFSSQTV 443  
QY 462 DSIPLFEMSVYNSEKSYDGVESKRMGMQLDRASSSLYVAESTCVIKVPLGRG--ERHG 519  
DB 446 DSVVIEBLQV.P-----PGVPKNLYVYRNDGDSKLVVVSDEILATKLRCGSKT 490  
QY 520 KOKKCTIASRDPYGMKEGACSHL-SPN---SRJTFQC-ERKNTDGLGCHNSFVAL 575  
DB 500 NCRB-CVSLQCPYCAMDNVELKCTAVGSPDMSAGKRRFIQN-SDGEHKACGGRPOTEIV- 557  
QY 576 NGRSSSLJPTTTTSDSA-----QGVESRGNLMDWKHLSDPSTDFLGAVSSHQK 628  
DB 558 -----ASVPFQPTKSGGDPVHS:HOAEFEPE---DNEVLVGDSKVPNTLABINHA 610  
QY 629 DKGVIRESYKXGHCQGVFV-----LAL-----AVLAFYNGAVFS 666  
DB 611 GSK-----JFSQSKLPYVTAETITIAVTSCLGALVGFISGFLPS 692

RESULT 11  
US-08-121-713D-60  
Sequence 60, Application US/68121713D

Patent No. 5639856  
GENERAL INFORMATION:  
APPLICANT: Goodman, Corey S.  
APPLICANT: Kolodkin, Alex L.  
APPLICANT: Matthes, David  
APPLICANT: Bentley, David R.  
APPLICANT: O'Connor, Timothy  
TITLE OF INVENTION: The Semaphorin Gene Family  
NUMBER OF SEQUENCES: 100  
CORRESPONDENCE ADDRESS:  
ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 Bush Street, Suite 3200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,713D

Query Match 15.18; Score 880; DB 1; Length 650;  
Best Local Similarity 35.84; Proc. No. 1.3e-74;  
Matches 207; Conservative 95; Mismatches 194; Indels 82; Gaps 19;  
QY 114 KQCHNFIKVLKXNDALFVCGTNAPNSCRNYKMD-----TLEPPCEBFSGMARCPYDA 169  
DB 1 EDCQNVIRIWPSPGSLFVCGTNSFRPNCNTYIISDSKYTLEA---TKXGQAVCPYDP 57  
QY 170 KXKVALFADCKLYSATVTFPLAIDAVIYRSLGESPTRTVKEDSKWLKEPYVQAVGYG 229  
DB 58 RHNTSVJAONELYSGVAVDFSGSDPIIYRE-----PQTEQYDLSLSLNAPNFVSSTQG 112  
QY 230 DYIYFFPREZAVENYTKGVVFEPRVACVCKNDMGGSQVLEKOWTFPKAHLNCSVPGDS 289  
DB 113 DFVYFFPREZAVENYTKGVVFEPRVACVCKNDMGGSQVLEKOWTFPKAHLNCSVPGDS 289  
QY 290 HFYFNILQAVTDVIR-----INGREWLIAETSTVPNSIPGSAVCAVDMGLIASVFTGRKE 345  
DB 172 PFYFNEIQSASNLVEGOVSGMSSKLIYGVFTPNSIPGSAVCAVCAFAAQD-ADTFEGQFKE 231  
QY 346 QKSPDSQVTPVDEBVPKPRPGCCAGSSSJERYATSNRFPDCTIANFMTPLMDPRAVPSIFNRPM 410  
DB 232 QTG-NSNMLPYNNAKVDPDRPGSC-----HNDSRALPDP-LNFIKTHSLMDENYPAF 283  
QY 406 FNRWFLRTKVRVELTKIAYD---TAAGPYQNHVTVFLGSEKGLILKFIARIGNSGFL 460  
DB 284 FSCPILVRTSTIYFTQAVDAQIKTPG--KYCVIFVGTGDKGIKISVNASADSADK 341  
QY 461 NCSLFEEMSVYNSEKSYDGVESKRMGMQLDRASSSLYVAESTCVIKVPLGRG--ERHG 519  
DB 342 VTSVVEIEDVLTQS-----SPIRNLEIVRTMQVCPKDGSDYDCKLIIVTDSQVA-OLH 397  
QY 514 RC--ERHCKCKTKTIASRDPYGMKEGACSHL-SPN---SRJTFQC-ERKNTDGLGCHNSFVAL 575  
DB 396 RCHNDKLTSCSE-CVALCQPYCAMDNVELKCTAVGSPDMSAGKRRFIQN-SDGEHKACGGRPOTEIV- 557  
QY 569 HNSEFVNLNGHSSSLJPTTTTSDSAQGVESRGNLMDWKHLSDPSTDFLGAVSSHQK 628  
DB 450 -----HACPSGKINSKCANAGECKGFRNDM-----DLIDS-----RRQ 483  
QY 629 DKGVIRESYKXGHCQGVFV-----LAL-----AVLAFYNGAVFS 666  
DB 484 SKQOEIIDKDKNFDIINAQYTVETLWMAVLASGIFS 521

RESULT 12  
US-08-935-248-60  
Sequence 60, Application US/24835248  
Patent No. 5927826  
GENERAL INFORMATION:  
APPLICANT: Goodman, Corey S.  
APPLICANT: Kolodkin, Alex L.  
APPLICANT: Matthes, David  
APPLICANT: Bentley, David R.  
APPLICANT: O'Connor, Timothy

## TITLE OF INVENTION: The Semaphorin Gene Family

NUMBER OF SEQUENCES: 100  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 Bush Street, Suite 3200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/615,266  
FILING DATE:

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/121,713  
FILING DATE: 13-SEP-1993

## ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: B94-002-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415)343-4342  
TELEFAX: (415) 343-4342  
TELEX:

## INFORMATION FOR SEQ ID NO: 60:

SEQUENCE CHARACTERISTICS:

LENGTH: 650 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-835-268-60

Query Match 16.1%; Score 880; DB 1; Length 650;

Best Local Similarity 35.8%; Pred. No. 1.3e-74; Indels 82; Gaps 19;  
Matches 207; Conservative 95; Mismatches 194;

QY 114 KDCRPFKVLKKNDALFVCGTNAFNSCRNYKMD---TLEPFGEFSGMARCPYDA 169  
DB 1 EDCQYIRIMVPSGRLEFVCGTNSFRMNCNYIISDSNYTLEA---TKNGQAVCPYP 57  
QY 170 KHANVALPAGKLYSATVDTFLAIDAVIYKSLGESPTRTVKHDSKWLKEPYTVQAVY 229  
DB 58 RHNSTSVLADNELYSCTVADFGSDPIIYRE----PLQTEQYDSLSLNAPNFVSFTG 112  
QY 230 DYIYFFPRETAVENTMGKTVVFRVAQVCKNDXGSGORVLEKQWTSFLKARLNCSPGDS 289  
DB 113 DFVYFFPRETAVENTMGKTVVFRVAQVCKNDXGSGORVLEKQWTSFLKARLNCSPGDS 171  
QY 290 HFYFNILQAVTDVIR-----INGRDVWLATFTPTNSIPGSAVCAYDMLDIASVFTGRFKE 345  
DB 172 PFYFNEQSASNLVEGOYGSMSKLVGVNTFSNSIPGSAVCAPALQDIADTFEQRPKE 231  
QY 346 QKSFDSWTVPDERVPKPRFGCAGSSSLERATSNRPDITLNKTHPLVDENAFSI 405  
DB 232 QTSINSXWLPVNAKVPDPFRGSC-----HNSGRALPDTLNFKTHSLXCDENYAP 283  
QY 406 FNRWEVTRVYRLTRIADV---TAAGPYQKHVVVLGSEKGIILKFL-ARIGKSGFL 460  
DB 284 FSPQILVTRSTVIRFTQIADIAKIKTEGG--KTYDVIFVUTDGHGKIKNVNESASASX 341  
QY 461 NCSLFEEMSYKSEKSGYDGVQDKRM-QNQLDRASSSLY-----VAFSTCVIKVPLG 513  
DB 342 VTSVVIIEIDVLEKNS-----EPINKLEVRNYQVQDKGSGYDGGKILITDSQWALQW 397  
QY 514 RC--ERAGKCKKTCAGSDPYCQWKEGGAC-SHLSFN--SPITFQCIERKNTGSLGDC 569  
DB 398 RCNCKKITSGSE-CVAIQEYCAWCKIAGKCRSGFAPRMDEENTFYCNVATGQ-----449

## RESULT 13

US-09-060-692-60  
Sequence 60: Application US/09060632  
Patent No. 5335965

## GENERAL INFORMATION:

APPLICANT: Goodman, Corey S.  
APPLICANT: Koldkin, Alex L.  
APPLICANT: Matthes, David  
APPLICANT: Bentley, David R.  
APPLICANT: O'Connor, Timothy

## TITLE OF INVENTION: The Semaphorin Gene Family

NUMBER OF SEQUENCES: 100

## CORRESPONDENCE ADDRESS:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 Bush Street, Suite 3200  
CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94104

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/060,692

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/121,713

FILING DATE: 13-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A.

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: B94-002-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415)343-4341

TELEFAX: (415) 343-4342

TELEX:

INFORMATION FOR SEQ ID NO: 60:

SEQUENCE CHARACTERISTICS:

LENGTH: 650 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-060-692-60

Query Match 16.1%; Score 880; DB 2; Length 650;

Best Local Similarity 35.8%; Pred. No. 1.3e-74; Indels 82; Gaps 19;  
Matches 207; Conservative 95; Mismatches 194;

QY 114 KDCRPFKVLKKNDALFVCGTNAFNSCRNYKMD---TLEPFGEFSGMARCPYDA 169  
DB 1 EDCQYIRIMVPSGRLEFVCGTNSFRMNCNYIISDSNYTLEA---TKNGQAVCPYP 57  
QY 170 KHANVALPAGKLYSATVDTFLAIDAVIYKSLGESPTRTVKHDSKWLKEPYTVQAVY 229  
DB 58 RHNSTSVLADNELYSCTVADFGSDPIIYRE----PLQTEQYDSLSLNAPNFVSFTG 112  
QY 230 DYIYFFPRETAVENTMGKTVVFRVAQVCKNDXGSGORVLEKQWTSFLKARLNCSPGDS 289  
DB 113 DFVYFFPRETAVENTMGKTVVFRVAQVCKNDXGSGORVLEKQWTSFLKARLNCSPGDS 171  
QY 290 HFYFNILQAVTDVIR-----INGRDVWLATFTPTNSIPGSAVCAYDMLDIASVFTGRFKE 345





```

:
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 58/835,268
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman, Richard A.
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: B94-002-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415)343-4341
: TELEFAX: (415) 343-4342
: TELEX:
:
: INFORMATION FOR SEQ ID NO: 60:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 650 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-060-610-60
:
: Query Match 16.1%; Score 860; DB 4; Length 650;
: Best Local Similarity 35.6%; Pred. No. 1,3e 74;
: Matches 207; Conservative 95; Mismatches 194; Indels 82; Gaps 13;
:
: QY 1-4 KPECHNFKVLLKKNDDALFVGSTNAENPSQRYQND-----TLPEFGDEFSMAQCPYDA 169
: DDB 1 EDDCQNYRIMVVPSPGRLFVGSTNSFRPMCTYIISDSNYTLEA---TKGQAVCPYDP 57
:
: QY 170 KIANVALFADGKLYANVTTFEIALDAVIYRSGESPTLETYKHDGKMKKEPYFQAVDYG 229
: DDB 56 RHNS'SVLADNELYSGTVADPSGDFIIVAE-----PLQTQYDLSLSNAPNFUSSTQD 112
:
: QY 230 DYIYFFPBIAYEYNTMGKVPFPAQVCKDKDQSGQSVLEKQNTSPFKARINSVFQDS 289
: DDB 113 DEVFYFPECTAVEFTGCKAIVRYARVCKWKGQPHR-FENRMTSPFKASLNCSTPGDY 171
:
: QY 290 HYPFNILCAVTDVIR-----LNGHGVVLATESPVYSIFGGAVGAYDMADIASVFTGRKE 345
: DDB 172 PFFYFNEIQASNLVGGVGSMSKLYGVNTPSHSIFGSNVCATFALQDADTFEGQKE 231
:
: QY 346 QKSPSTWTPVLEVPKPPRPGCCAGSSSLERYATNRPDITLNFKTHPIMBEAVTSI 405
: DDB 232 QTGKNSMIFVNNAKVPPDPGSG-----HNSRALPQPTLNPKTHSLYDENVAF 283
:
: QY 406 ENRPWELRTMAYRLTKIAD-----TAAGPYQHTVYFLGSEKGIILKFL-ARIONSGEL 460
: DDB 284 FSPQ-LVRTSTIYRFTQIADACIKTPGG--KYDYVTFVGTGHHGKIISKVAESACSAK 341
:
: QY 461 NDSJFLEMSVYNSEKSYDGVDEKRLX-GMQLDRASSSLY-----VAFSTCVIKVPLG 513
: DDB 342 VTSVVEIEDVLTKS-----EPINLEIVRTNOYDCKDGSYDGGKLIIVTDSQVVALCLH 397
:
: QY 514 RC--ERHGKCKTCTASRDYPGMIKEGGAC-SHLSPN--SELTPEQDIERGNTDGLGDC 569
: DDB 398 RCHNDKITSCE-SVALQDPYCAWQKAGKCRSHGAPRLBENFYQKAVATGQ----- 449
:
: QY 569 HNSFVALNCHSSSLIPSTTSDSTAGEYBSRGMLDWKHLIDSPDSTDPLGAVSSHQ 626
: DDB 450 -----HAACPSGKINSKANAGECKGRNDN-----DLLES-----PRQ 483
:
: QY 629 DKKGVIRESYLKGHCLVPTLLAIATVAFVMGAVES 666
: DDB 484 SKDQEIIONDKNFEDIINAQYTVETLVAVLAGSIPS 521

```

Search completed: October 23, 2003, 17:10:59

Job time : 23 secs





XX OS Homo sapiens.  
XX PN EP1074617-A2.  
XX PD 07-FEB-2001.  
XX PF 28-JUL-2000; 2000EP-0116126.  
XX PR 29-JUL-1999; 99JP-0248036.  
XX PR 27-AUG-1999; 99JP-0300253.  
XX PR 11-JAN-2000; 2000JP-0118776.  
XX PR 02-MAY-2000; 2000JP-0183767.  
XX PR 09-JUN-2000; 2000JP-0241899.  
XX PA (HELI-) HELIX RPS INST.  
XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI: 2001-318749/34.  
XX PR Primer sets for synthesizing polynucleotides, particularly the 5602  
XX full-length cDNAs defined in the specification, and for the detection  
XX and/or diagnosis of the abnormality of the proteins encoded by the  
XX full-length cDNAs.  
XX Claim 8: SEQ ID NO: 2537p - CD ROM; English.  
XX The present invention describes primer sets for synthesizing 5602  
XX full-length cDNAs defined in the specification, where a primer set  
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
XX to the complementary strand of a polynucleotide which comprises one of  
XX the 5602 nucleotide sequences defined in the specification, where the  
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
XX of an oligonucleotide comprising a sequence complementary to the  
XX complementary strand of a polynucleotide which comprises a 5'-end  
XX sequence and an oligonucleotide comprising a sequence complementary to a  
XX polynucleotide which comprises a 3'-end sequence, where the  
XX oligonucleotide comprises at least 15 nucleotides and the combination of  
XX the 5'-end sequence/3'-end sequence is selected from those defined in  
XX the specification. The primer sets can be used in antisense therapy and  
XX in gene therapy. The primers are useful for synthesizing polynucleotides,  
XX particularly full-length cDNAs. The primers are also useful for the  
XX detection and/or diagnosis of the abnormality of the proteins encoded by  
XX the full-length cDNAs. The primers allow obtaining of the full-length  
XX cDNAs easily without any specialized methods. AAH03156 to AAH3628 and  
XX AAH13633 to AAH18742 represent human cDNA sequences; AAH2446 to  
XX AAH95883 represent human amino acid sequences; and AAH13629 to AAH13642  
XX represent oligonucleotides, all of which are used in the exemplification  
XX of the present invention.  
XX Sequence 2123 BP; 550 A; 564 C; 528 G; 381 T; 0 other;  
Query Match 100.0%; Score 216; DB 22; Length 2123;  
Best Local Similarity 100.0%; Pred. No. 3.5e-52;  
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCGCGCGCGCGCGCGAGGGTGGACTCCAGCTGCACAGCTCCAGCGCATCTGGC 60  
DB 1432 CCGCGCGCGCGCGCGAGGGTGGACTCCAGCTGCACAGCTCCAGCGCATCTGGC 1491  
QY 61 CAGGCGGTGACTGTCTCGAGGAGCGCCAGCGCTCAACGCTCAACTCACTGACAGGTGG 120  
DB 1492 CAGGCGGTGACTGTCTCGAGGAGCGCCAGCGCTCAACGCTCAACTCACTGACAGGTGG 1551  
QY 121 GGGCTGAAGCGTAGCGGCTCGCTAAAGCGGAGCGGTACCGCCCAACACCATCTTTCCTCC 180  
DB 1552 GGGCTGAAGCGTAGCGGCTCGCTAAAGCGGAGCGGTACCGCCCAACACCATCTTTCCTCC 1611  
QY 181 CTTTCCACATCCATGAGCCCAATGATGCGTGACA 216  
DB 1612 CTTTCCACATCCATGAGCCCAATGATGCGTGACA 1647

RESULT 4  
AAK94365  
ID AAK94365 standard; cDNA: 2227 BP.  
XX AC AAK94365.  
XX UF 06-NOV-2001 (first entry)  
XX DE Human full-length cDNA, SEQ ID NO: 3087.  
XX KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.  
XX OS Homo sapiens.  
XX PN EP140094-A2.  
XX PD 05-SEP-2001.  
XX PF 07-JUL-2000; 2000EP-0114069.  
XX PR 08-JUL-1999; 99JP-0194486.  
XX PR 11-JAN-2000; 2000JP-0118774.  
XX PR 02-MAY-2000; 2000JP-0183765.  
XX PA (HELI-) HELIX RPS INST.  
XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX WPI: 2001-524255/58.  
XX PR P-PS28; ANM93444.  
XX PR 830 Primers useful for synthesizing full length cDNA clones and their  
XX use in genetic manipulation.  
XX Claim 8: SEQ ID NO 3087; 1380pp - sequence listing; English.  
XX The invention relates to primers for synthesizing full length cDNA  
XX clones. 830 cDNA molecules encoding a human protein have been  
XX isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
XX molecules have been determined. Primers for synthesizing the full length  
XX cDNA are useful for clarifying the function of the protein encoded by  
XX the cDNA. The full length clones were obtained by construction of full  
XX length enriched cDNA libraries that were synthesised by the oligo-capping  
XX method. The primers enable the production of the full length cDNA easily  
XX without any special methods. The present sequence is a full length  
XX human cDNA of the invention.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in CD-ROM format directly from EPO.  
XX Sequence 2227 BP; 567 A; 700 C; 560 G; 400 T; 0 other;  
Query Match 100.0%; Score 216; DB 22; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 3.6e-52;  
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCGCGCGCGCGCGCGAGGGTGGACTCCAGCTGCACAGCTCCAGCGCATCTGGC 60  
DB 1526 CCGCGCGCGCGCGCGAGGGTGGACTCCAGCTGCACAGCTCCAGCGCATCTGGC 1585  
QY 61 CAGGCGGTGACTGTCTCGAGGAGCGCCAGCGCTCAACGCTCAACTCACTGACAGGTGG 120  
DB 1586 CAGGCGGTGACTGTCTCGAGGAGCGCCAGCGCTCAACGCTCAACTCACTGACAGGTGG 1645  
QY 121 GGGCTGAAGCGTAGCGGCTCGCTAAAGCGGAGCGGTACCGCCCAACACCATCTTTCCTCC 180  
DB 1646 GGGCTGAAGCGTAGCGGCTCGCTAAAGCGGAGCGGTACCGCCCAACACCATCTTTCCTCC 1705  
QY 181 CTTTCCACATCCATGAGCCCAATGATGCGTGACA 216  
DB 1706 CTTTCCACATCCATGAGCCCAATGATGCGTGACA 1741

Query Match 100.0%; Score 216; SS 24; Length 2262;  
Best Local Similarity 100.0%; Pred. No. 3,6e-52;

CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX Sequence 2293 BP; 579 A; 715 C; 531 G; 420 T; 0 other;

Query Match 100.0%; Score 216; DB 22; Length 2293;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-52;  
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCCGCCGCCGCCGAGAGGTGGATCTCATCCAGCTGACACACTCCAGCCATGTCG 60  
 DB 1508 CCGCCGCCGCCGCCGAGAGGTGGATCTCATCCAGCTGACACACTCCAGCCATGTCG 60  
 QY 61 CAGCCCTGACTGTCGAGGACGCCGAGGCTGACGCTGACACTCAGTGCACAGGTCG 120  
 DB 1668 CAGCCCTGACTGTCGAGGACGCCGAGGCTGACGCTGACACTCAGTGCACAGGTCG 120  
 QY 121 GGGCTGAAGCTAGCGCTGCTGCTAAACCGGAGTACAGCTGACAGCTCTTGTGTCG 180  
 DB 1728 GGGCTGAAGCTAGCGCTGCTGCTAAACCGGAGTACAGCTGACAGCTCTTGTGTCG 180  
 QY 181 CTTCACATCATGATGAGCCCAATGATGCGTGATCA 216  
 DB 1788 CTTCACATCATGATGAGCCCAATGATGCGTGATCA 216

# RESULT 7

AAH17861  
 ID AAH17861 standard; cDNA; 2346 BP.  
 AC AAH17861:  
 XX  
 XX  
 XX 26-JUN-2001 (first entry)  
 XX Human cDNA sequence SEQ ID NO:17567.  
 DE  
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX Homo sapiens.  
 XX  
 XX EP107467-A2.  
 XX  
 XX 07-FEB-2001.  
 XX  
 XX 28-JUN-2000; 2000EP-011612.  
 XX  
 XX 29-JUL-1999; 99JP-C248036.  
 PR 27-AUG-1999; 99JP-C300253.  
 PR 11-JAN-2000; 2000JP-C118776.  
 PR 02-MAY-2000; 2000JP-C0183757.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto C;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 XX WPI: 2001-318749/34.  
 DR  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 XX Claim 8; SEQ ID 17567; 2537pp - CD ROM; English.  
 PS  
 XX The present invention describes primer sets for synthesizing 5602  
 CC

CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesising polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX Sequence 2306 BP; 579 A; 739 C; 577 G; 411 T; 0 other;

Query Match 100.0%; Score 216; DB 22; Length 2306;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-52;  
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCCGCCGCCGCCGAGAGGTGGATCTCATCCAGCTGACAGCTCTTGTGTCG 60  
 DB 1615 CCGCCGCCGCCGCCGAGAGGTGGATCTCATCCAGCTGACAGCTCTTGTGTCG 60  
 QY 61 CAGCCCTGACTGTCGAGGACGCCGAGGCTGACGCTGACACTCAGTGCACAGGTCG 120  
 DB 1675 CAGCCCTGACTGTCGAGGACGCCGAGGCTGACGCTGACACTCAGTGCACAGGTCG 120  
 QY 121 GGGCTGAAGCTAGCGCTGCTGCTAAACCGGAGTACAGCTGACAGCTCTTGTGTCG 180  
 DB 1735 GGGCTGAAGCTAGCGCTGCTGCTAAACCGGAGTACAGCTGACAGCTCTTGTGTCG 180  
 QY 181 CTTCACATCATGATGAGCCCAATGATGCGTGATCA 216  
 DB 1795 CTTCACATCATGATGAGCCCAATGATGCGTGATCA 216

# RESULT 8

AAH68253  
 ID AAH68253 standard; cDNA; 2592 BP.  
 XX  
 XX AAH68253:  
 AC  
 XX  
 XX 13-FEB-2002 (first entry)  
 XX  
 XX DNA encoding novel human diagnostic protein #4037.  
 DE  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 KW  
 XX Homo sapiens.  
 XX  
 XX WO200175067-A2.  
 XX  
 XX 11-OCT-2001.  
 XX  
 XX 30-MAR-2001; 2001WO-050831.  
 PF  
 XX 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 XX  
 XX Ormanac RT, Liu C, Tang YT;  
 XX

DR WPI: 2001-639362/73.  
 DR P-PSDB: ABG04066.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 1: SEQ ID No 4057; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 2592 BP; 655 A; 754 C; 659 G; 524 T; 0 other;  
 Query Match 100.0%; Score 216; DB 23; Length 2592;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-52;  
 Matches 216; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGCGCGCCGCCCGCAGAGAGGTGGACTCCATCCAGGTGCACAGCTCCAGCCATCTGGC 60  
 DB 2374 CCGCGCGCCGCCCGCAGAGAGGTGGACTCCATCCAGGTGCACAGCTCCAGCCATCTGGC 2433  
 QY 61 CAGCGCGTGACTGTCTGAGGACAGCCAGCTCCAGCTTACACCTACACCTACGACAGGTG 125  
 DB 2434 CAGCGCGTGACTGTCTGAGGACAGCCAGCTCCAGCTTACACCTACACCTACGACAGGTG 2493  
 QY 121 GGGCTGAAGCGTACGCCCTCGCTAAAGCGGAGTACGCCCAAGCATCTCTTGGTCTCG 180  
 DB 2494 GGGCTGAAGCGTACGCCCTCGCTAAAGCGGAGTACGCCCAAGCATCTCTTGGTCTCG 2553  
 QY 181 GTTCCACATCCATGAAGCGCAATGATCGGTGTACA 216  
 DB 2554 GTTCCACATCCATGAAGCGCAATGATCGGTGTACA 2580  
 RESULT 9  
 AAS68807  
 ID AAS68807 standard; cDNA: 3039 BP.  
 XX  
 AC AAS68807;  
 XX  
 DT 13-FEB-2002 (first entry)  
 DE DNA encoding novel human diagnostic protein #4611.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensics;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200275057-A2.  
 XX  
 DT 11-OCT-2001.  
 PD

30-MAR-2001; 2001WO-US86631.  
 31-MAR-2000; 2000US-0540217.  
 23-AUG-2000; 2000US-0649167.  
 (NYSE-) HVSFO INC.  
 Dr. Daniel R. Liu C, Tang Y.  
 WPI: 2001-639362/73.  
 P-PSDB: ABG04620.  
 New isolated polynucleotide and encoded polypeptides, useful in  
 diagnostics, forensics, gene mapping, identification of mutations  
 responsible for genetic disorders or other traits and to assess  
 biodiversity.  
 Claim 1: SEQ ID No 4611; 103pp; English.  
 The invention relates to isolated polynucleotide (I) and  
 polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 and gene mapping, and in recombinant production of (II). The  
 polynucleotides are also used in diagnostics as expressed sequence tags  
 for identifying expressed genes. (I) is useful in gene therapy techniques  
 to restore normal activity of (II) or to treat disease states involving  
 a food supplement. (II) and its binding partners are useful in medical  
 imaging of sites expressing (II). (I) and (II) are useful for treating  
 disorders involving aberrant protein expression or biological activity.  
 The polypeptide and polynucleotide sequences have applications in  
 diagnostics, forensics, gene mapping, identification of mutations  
 and to produce other types of data and products dependent on DNA and  
 amino acid sequences. AAS64197-AAS94564 represent novel human  
 diagnostic coding sequences of the invention.  
 Note: The sequence data for this patent did not appear in the printed  
 specification, but was obtained in electronic format directly from WIPO  
 at ftp.wipo.int/pub/published\_pct\_sequences.  
 Sequence 3039 BP; 741 A; 869 C; 781 G; 648 T; 0 other;  
 Query Match 100.0%; Score 216; DB 23; Length 3039;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-52;  
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGCGCGCCGCCCGCAGAGAGGTGGACTCCATCCAGGTGCACAGCTCCAGCCATCTGGC 60  
 DB 2821 CGCGCGCCGCCCGCAGAGAGGTGGACTCCATCCAGGTGCACAGCTCCAGCCATCTGGC 2880  
 QY 61 CAGCGCGTGACTGTCTGAGGACAGCCAGCTCCAGCTTACACCTACACCTACGACAGGTG 120  
 DB 2881 CAGCGCGTGACTGTCTGAGGACAGCCAGCTCCAGCTTACACCTACACCTACGACAGGTG 2840  
 QY 121 GGGCTGAAGCGTACGCCCTCGCTAAAGCGGAGTACGCCCAAGCATCTCTTGGTCTCG 180  
 DB 2941 GGGCTGAAGCGTACGCCCTCGCTAAAGCGGAGTACGCCCAAGCATCTCTTGGTCTCG 3000  
 QY 181 GTTCCACATCCATGAAGCGCAATGATCGGTGTACA 216  
 DB 3001 GTTCCACATCCATGAAGCGCAATGATCGGTGTACA 3035  
 RESULT 10  
 AAS99721  
 ID AAS99721 standard; cDNA: 3039 BP.  
 XX  
 AC AAS99721;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX

DE DNA encoding novel human diagnostic protein #25525.

XX Human: chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX Homo sapiens.

XX WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2002; 2001US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX P2 Drmanac RT, Liu C, Tang YJ;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG25534.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX Claim 1; SEQ ID NO 25525; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (II) is useful in gene therapy techniques  
CC to restore normal activity of (I) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AA564197-AA594564 represent novel human  
CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pat\_sequences.

XX Sequence 3039 BP; 741 A; 865 C; 781 G; 642 T; 0 other;

XX Query Match 100.0%; Score 216; DB 23; Length 3039;  
XX Best Local Similarity 100.0%; Pred. No. 3.8e-52;  
XX Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCCGCCGCCGACAGAGGTGGACTCCAGCTGCACAGCTCCGACGACATCGGC 60

DB 2821 CGCGCCGCCGCCGACAGAGGTGGACTCCAGCTGCACAGCTCCGACGACATCGGC 2800

QY 61 CAGGCCGTGACTGTCTCGAGCAGCCGCCAGCTCAACGCTCAACTCAGTCAGCAGGTGG 120

DB 2881 CAGGCCGTGACTGTCTCGAGCAGCCGCCAGCTCAACGCTCAACTCAGTCAGCAGGTGG 2340

QY 121 GGGCTGAGAGCTACCCCTCGCTAAGCCGAGGAGTACCCGCCAAGACCATCTTTCCTCC 180

DB 2941 GGGCTGAGAGCTACCCCTCGCTAAGCCGAGGAGTACCCGCCAAGACCATCTTTCCTCC 3000

QY 18: GTTCCACATCCATGAGCCCAATCATCGCTGATCA 216

DB 3001 GTTCCACATCCATGAGCCCAATCATCGCTGATCA 3036

RESULT 11

AAH17625

ID AAH17625 standard; cDNA; 3041 BP.

XX AC AAH17625;

XX DT 24-JUN-2001 (first entry)

XX DE Human cDNA sequence SEQ ID NO:17153.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PK EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI) HEMEX RES INST.

XX PI Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama I, Wakanatsu A, Naqai K, Otsuki I;

XX DR WPI; 2001-318749/34.

XX PR Primer sets for synthesizing polynucleotides, particularly the 5602

CC full-length cDNAs defined in the specification, and for the detection

CC and/or diagnosis of the abnormality of the proteins encoded by the

CC full-length cDNAs.

XX Claim 8; SEQ ID 17153; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesising polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to

CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

XX Sequence 3041 BP; 804 A; 886 C; 714 G; 637 T; 0 other;

XX Query Match 100.0%; Score 216; DB 22; Length 3041;

XX Best Local Similarity 100.0%; Pred. No. 3.8e-52;

XX Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCCGCCGCCGACAGAGGTGGACTCCAGCTGCACAGCTCCGACGACATCTGGC 60

|||||







[illegible]

## ALIGNMENTS











double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p773 vector (Pharmacia). Library constructed by Bento Soares and M. Patricia Bonaldi. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino.

BASE COUNT 111 a 167 c 163 g 77 t 4 others  
ORIGIN

Query Match 94.1%; Score 212; DP 14; Length 442;  
Best Local Similarity 94.1%; Pred. No. 1, 9e-44;  
Matches 212; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGCGCCGCGCGCGGAGA GATGGACTCCATCCAGGTCACAGGTCGACGATGGG 60  
D5 239 CGCGCCGCGCGCGGAGA GATGGACTCCATCCAGGTCACAGGTCGACGATGGG 237  
QY 61 CAGCGCTGACTGTCTGAGAGCGCCAGCCTCAAGCTTAAAGCTGACGACAGGTCG 126  
DB 298 CAGCGCTGACTGTCTGAGAGCGCCAGCCTCAAGCTTAAAGCTGACGACAGGTCG 357  
QY 121 GGGCTGAAGCGTACGGCTCTGCTAAAGCGGAGCTACCGCCGCAAGATCCTTTGCTAC 186  
DB 358 GGGCTGAAGCGTACGGCTCTGCTAAAGCGGAGCTACCGCCGCAAGATCCTTTGCTAC 417  
QY 181 CTTTCCATCCATGAGAGCGCAATGATGGGCTGACA 216  
DB 418 CTTTCCATCCATGAGAGCGCAATGATGGGCTGACA 453

RESULT 10  
N46020  
LOCUS 410 bp mRNA linear EST 14-PEB-1996  
DEFINITION IMAGE:273200 5', mRNA sequence.  
ACCESSION N46020  
VERSION N46020.1 GI:1187186  
KEYWORDS EST.  
SOURCE Homo sapiens (human);  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 410)  
AUTHORS Hillier, L., Clark, R., Dubouque, T., Ellison, K., Hawkins, M., Hultman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Maric, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlschlag, P. and Wilson, R.  
TITLE The WashU-Merck EST Project  
JOURNAL Unpublished  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8502, St. Louis, MO 63109  
TEL: 314 286 1800  
FAX: 314 286 1810  
Email: est@wustl.edu  
This clone is available royalty-free through IMAGE; contact the IMAGE Consortium (image@imgl.gov) for further information.  
Seq primer: T7  
High quality sequence stop: 329.

FEATURES  
source  
1..410  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3882842"  
/db\_xref="taxon:9606"  
/clone="IMAGE:273200"  
/sex="Male"  
/tissue\_type="melanocyte"  
/lab\_host="DHCH (ampicillin resistant)"  
/clone\_lib="Soares melanocyte 2NBM"  
/note="Vector: p7730 (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5'

TCTTACCAATCTGAAGTGGAGCGCGCGAGTCTTTTCTTTTCTTTT 31.  
double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p773 vector (Pharmacia). Library constructed by Bento Soares and M. Patricia Bonaldi. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino.

BASE COUNT 99 a 143 c 161 g 62 t 5 others  
ORIGIN

Query Match 94.4%; Score 203.6; DP 14; Length 410;  
Best Local Similarity 94.4%; Pred. No. 1, 9e-44;  
Matches 206; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 CGCGCCGCGCGCGGAGA GATGGACTCCATCCAGGTCACAGGTCGACGATGGG 66  
D5 73 CGCGCCGCGCGCGGAGA GATGGACTCCATCCAGGTCACAGGTCGACGATGGG 137  
QY 67 GTGACTGTCTGAGAGCGCCAGCCTCAAGCTTAAAGCTGACGACAGGTCG 126  
DB 138 GTGACTGTCTGAGAGCGCCAGCCTCAAGCTTAAAGCTGACGACAGGTCG 197  
QY 127 AAGGTAGCGCTGCTGCTAAAGCGGAGCTACCGCCGCAAGATCCTTTGCTAC 186  
DB 198 AAGGTAGCGCTGCTGCTAAAGCGGAGCTACCGCCGCAAGATCCTTTGCTAC 257  
QY 187 ACATCCATGAAGCGCAATGATGGGCTGACA 216  
D5 258 ACATCCATGAAGCGCAATGATGGGCTGACA 287

RESULT 11  
CB577257  
LOCUS 632 bp mRNA linear EST 03-APR-2003  
DEFINITION AMENUC:CDRGL-00015-G10-A cdrgl (10898) Rattus norvegicus CDNA clone cdrgl-00015-g10 5', mRNA sequence.

ACCESSION CB577257  
VERSION CB577257  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 632)  
AUTHORS Amgen EST Program  
TITLE Amgen Rat EST Program  
JOURNAL Unpublished  
COMMENT Contact: Dan Fitzpatrick  
Amgen, Inc  
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA  
TEL: 805 447-4881  
Plate: 00015 row: g column: 10.

FEATURES  
Location/Qualifiers  
1..632  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="cdrgl-00015-g10"  
/tissue\_type="Chung Model Ipsilate"  
/clone\_lib="cdrgl (10898)"  
/note="Vector: pSPORT1; Chung Model Rat DRG Left L5/L6"

BASE COUNT 155 a 213 c 151 g 113 t  
ORIGIN

Query Match 89.6%; Score 193.6; DP 14; Length 632;  
Best Local Similarity 93.5%; Pred. No. 1, 1e-41;  
Matches 202; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 CGCGCGCGCGCGCGGAGGTCGACGTCGATCCATCCAGTGCACAGTCCTCCAGGATCTGGC 60  
DB 20 CCAGCTCCGCGCGCGGAGGTCGACGTCGATCCATCCAGTGCACAGTCCTCCAGGATCTGGC 79











Db 343 CTTGGAGCGGAGCACATCTTGGGCGCATGGCTTGGCGGGCGAGGTGGCGGG 284  
QY 62 AGCGCGTACTGTCAGGAGAGCGGAGCGCTCAAGCGCTCACTCACTCAAGGTCGG 122  
Db 283 CGTGGCGCGGAGCGGAGCGGCGGCGGCGGAGCGGAGCGGAGCGGAGCGG 224  
QY 122 GGCTGAAGCGTACCGCTCGCTAAAGCGGAGCTATCGCTCAAGCACTGCTGGCGGG 182  
Db 223 CCAGCCAGCGTTTACGCTGCTATGCTGTCAGCGGCGGAGCGGAGCGGAGCGG 164  
QY 182 T 182  
Db 163 T 163

## RESULT 8

US-09-252-991A-1737  
; Sequence 1737, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074.788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094.190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 1737  
; LENGTH: 1026  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-1737

Query Match 15.6% Score 33.8; DB 4; Length 1026;  
Best Local Similarity 54.4% Pred. No. 0.76;  
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;  
QY 62 AGCGCGTACTGTCAGGAGAGCGGAGCGCTCAAGCGCTCACTCACTCAAGGTCGG 122  
Db 271 AGCGGATCAACTCAAGGAGTCTCGAGGAGGCGGAGGAGTCTCGAGGTCGA 330  
QY 122 GGCTGAAGCGTACCGCTCGCTAAAGCGGAGCTATCGCTCAAGCACTGCTGGCGGG 182  
Db 331 GAAGGAAGAGCGGCGGAGCGGAGCGGCGGAGCGGAGCGGAGCGGAGCGG 390  
QY 182 TTTC 186  
Db 391 TTAC 395

## RESULT 9

US-09-252-991A-9254  
; Sequence 9254, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074.788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094.190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 9254  
; LENGTH: 2667

; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-9254  
Query Match 15.6% Score 33.8; DB 4; Length 2667;  
Best Local Similarity 49.2% Pred. No. 0.96;  
Matches 89; Conservative 0; Mismatches 92; Indels 0; Gaps 0;  
QY 2 CGCGCGCGCGCGCGGAGCGGTCATCTCACTCAAGTGCAGAGCTCCAGGCACTTGGCC 61  
Db 36 CGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 97  
QY 62 AGCGCGTACTGTCAGGAGAGCGGAGCGCTCAAGCGCTCACTCACTCAAGGTCGG 121  
Db 96 CGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 157  
QY 122 GGCTGAAGCGTACCGCTCGCTAAAGCGGAGCTATCGCTCAAGCACTGCTGGCGGG 181  
Db 158 CCAGCCAGCGTTTACGCTGCTATGCTGTCAGCGGCGGAGCGGAGCGGAGCGG 217  
QY 182 T 182  
Db 213 T 213

## RESULT 10

US-09-252-991A-1816  
; Sequence 1816, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074.788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094.190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 1816  
; LENGTH: 3228  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-1816

Query Match 15.6% Score 33.8; DB 4; Length 3228;  
Best Local Similarity 54.4% Pred. No. 1;  
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;  
QY 62 AGCGCGTACTGTCAGGAGAGCGGAGCGCTCAAGCGCTCACTCACTCAAGGTCGG 121  
Db 303 AGCGCGGATCAACTCAAGGAGTCTCGAGGAGGCGGAGGAGGAGGAGGAGGAGG 362  
QY 122 GGCTGAAGCGTACCGCTCGCTAAAGCGGAGCTATCGCTCAAGCACTGCTGGCGGG 181  
Db 363 GAAGGAAGAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 422  
QY 182 TTTC 186  
Db 423 TTAC 427

## RESULT 11

US-09-252-991A-2143/c  
; Sequence 2143, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136

```
? CURRENT APPLICATION NUMBER: US/09/252,991A
? PRIOR FILING DATE: 1998-02-18
? PRIOR APPLICATION NUMBER: US 60/074,768
? PRIOR FILING DATE: 1998-02-18
? PRIOR APPLICATION NUMBER: US 60/094,190
? PRIOR FILING DATE: 1998-07-27
? NUMBER OF SEQ ID NOS: 33142
? SEQ ID NO 2143
? LENGTH: 3546
? TYPE: DNA
? ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2143

Query Match
Best Local Similarity 15.4%; Score 33.8; DB 4; Length 3546;
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 62 AGCGCGTACTCTCTGATGAGCGCCAGCTTCAAGCGCTTACAACTCACTGACAGAGTGG 121
DB 3064 AGCGCGGATCAACAATCAAGTCTGAGCGAGGCGAGGAGTCACTGCTCCAGTGG 3005
QY 122 GCGTCAAGCGTACGCCCTTCTTAAGCGCGGAGTACGCGCCCAAGCATCTTTCGTC 181
DB 3004 GAAGAAGAGCGCGGCAAAAGGCGCGCCCTGACGACCTTATCAAGCTGGAGGCG 2945
QY 182 TTCC 186
DB 2944 TTAC 2940

RESULT 12
US-09-252-991A-7012/c
? Sequence 7012, Application US/09252991A
? Patent No. 6551795
? GENERAL INFORMATION:
? APPLICANT: Marc J. Rubenfield et al.
? TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
? FILE REFERENCE: 107196.176
? CURRENT APPLICATION NUMBER: US/09/252,991A
? PRIOR FILING DATE: 1998-02-18
? PRIOR APPLICATION NUMBER: US 60/074,768
? PRIOR FILING DATE: 1998-02-18
? PRIOR APPLICATION NUMBER: US 60/094,190
? PRIOR FILING DATE: 1998-07-27
? NUMBER OF SEQ ID NOS: 33142
? SEQ ID NO 7012
? LENGTH: 630
? TYPE: DNA
? ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7012

Query Match
Best Local Similarity 15.4%; Score 33.2; DB 4; Length 630;
Matches 71; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 7 CCGCGCCCGGAGAGGTGGAGTCCATCCAGTGCACAGTCCAGCTGCGAGAGG 66
DB 203 CCGCGCGGATCGAGTGGATCCGCGCATCGAGGATCGCGCGGCTGGCTGG 144
QY 67 GTGACTGTCTCGAGCGAGTCCAGTGTCAAGCGCTTACAACTCACTGACAAAGTCTGG 126
DB 143 CATCTATGTTGGAGCGCGGCTCCATGAGCAIGTTCACGCGCTTCCCAAGGAGGGGAT 64
QY 127 AGCGTAGGCCCTC 140
DB 83 AGCGTAGGCCCTC 70

RESULT 13
US-09-252-991A-7061
? Sequence 7061, Application US/09252991A
? Patent No. 6551795
? GENERAL INFORMATION:
? APPLICANT: Marc J. Rubenfield et al.
? TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
? FILE REFERENCE: 107196.136
? CURRENT APPLICATION NUMBER: US/09/252,991A
? PRIOR FILING DATE: 1998-02-18
? PRIOR APPLICATION NUMBER: US 60/074,768
? PRIOR FILING DATE: 1998-02-18
? PRIOR APPLICATION NUMBER: US 60/094,190
? PRIOR FILING DATE: 1998-07-27
? NUMBER OF SEQ ID NOS: 33142
? SEQ ID NO 7061
? LENGTH: 1548
? TYPE: DNA
? ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7061

Query Match
Best Local Similarity 15.4%; Score 33.2; DB 4; Length 1548;
Matches 71; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 7 CCGCGCCCGGAGAGGTGGAGTCCATCCAGTGCACAGTCCAGCTGCGAGAGG 66
DB 623 CCGCGCGGATCGAGTGGATCGCGCGATCGCGCGGCTGGCTGGCTGGCTGG 682
QY 67 GTGACTGTCTCGAGCGAGCGCGAGCTTCAAGCGCTTACAACTCACTGACAAAGTCTGG 126
DB 683 CATCTATGTTGGAGCGCGCGGCTCCATGAGCATGCTTCCAGCGCGTTCACAGGAGGCG 742
QY 127 AGCGTAGGCCCTC 140
DB 743 AGCGTAGGCCCTC 756
```



```

RESULT 15
US-08-993-359-29
; Sequence 29, Application US/58693359A
; Patent No. 6039942
; GENERAL INFORMATION:
; APPLICANT: Lassen, Soren F.
; APPLICANT: Bech, Lisbeth
; APPLICANT: Ohmann, Anders
; APPLICANT: Reichardt, Jens
; APPLICANT: Fuglsang, Claus C.
; APPLICANT: Ostergaard, Peter F.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 5363-500-15
; CURRENT APPLICATION NUMBER: 49/08/993-359A
; CURRENT FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 1480/96
; EARLIER FILING DATE: 1996-12-20
; EARLIER APPLICATION NUMBER: 1481/96
; EARLIER FILING DATE: 1996-12-20
; EARLIER APPLICATION NUMBER: 0301/97
; EARLIER FILING DATE: 1997-03-18
; EARLIER APPLICATION NUMBER: 0523/97
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 1386/97
; EARLIER FILING DATE: 1997-12-01
; EARLIER APPLICATION NUMBER: 60/046-082
; EARLIER FILING DATE: 1997-05-09
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 1536
; TYPE: DNA
; ORGANISM: Truncetes pubescens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79)...(1407)
; NAME/KEY: mat_peptide
; LOCATION: (130)...(1407)
; NAME/KEY: sig_peptide
; LOCATION: (79)...(129)
US-08-993-359-29

```

```

Query Match      15.3% Score 33; DB 3; Length 1536;
Best Local Similarity 53.5%; Pred. No. 1.4;
Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 11 GCGCGCAGGGTGGAGCTCATCCAGGIGACAGCTCCGAGCCATCTGGCGAGGGGTGA 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 950 CCGTGCAGGCGGTGGGTAATCAACGAGCTCATCGCGGCTCCACCGCGCAAGGTGT 1009
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 71 CTGTCTGAGSCAGCCGCTCAACGCTACAACTCACTGACAGGCTGGGGCTGAAGC 130
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1010 CCGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 131 GTAGCCCT 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1070 GCACCTCT 1078
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: September 30, 2003, 22:06:17  
Job time : 27.4914 secs

















QY 1 CCGCGCGCGCGCGCGAGGGTGGAGTCCATCCAGTGCACAGCTCCAGCATCTGGC 60  
 DB 2875 CCGCGCGCGCGCGCGAGGGTGGAGTCCATCCAGTGCACAGCTCCAGCATCTGGC 2894  
 QY 61 CAGCGCGTGAATGCTGAGCGACAGCGCTCAAGCCACAACTCACTGACAAAGTGG 120  
 DB 2895 CAGCGCGTGAATGCTGAGCGACAGCGCTCAAGCCACAACTCACTGACAAAGTGG 2954  
 QY 121 GGGCTGAAGCGTACGGCTGCTAAAGCGGAGCTACCGCCCAAGCACTCTTGGTCCG 240  
 DB 2955 GGGCTGAAGCGTACGGCTGCTAAAGCGGAGCTACCGCCCAAGCACTCTTGGTCCG 3114  
 QY 181 GTTCCATCATCATGAAGCCCAATGATGGGTGACA 216  
 DB 3015 GTTCCATCATCATGAAGCCCAATGATGGGTGACA 3050

RESULT 14  
 AK096337  
 LOCUS Homo sapiens cDNA FLJ39019 f.3, clone M249702594, highly similar to Homo sapiens semaphorin SEMA6A mRNA.  
 ACCESSION AK096337  
 VERSION 1  
 KEYWORDS cDNA; capping; fis (full insert sequencing).  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Niromiya K., Kagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H., Puriya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H., Yatsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K., Kaneko K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami S., Suzuki Y., Sugano S., Nishihara K., Masuko Y., Nagai T. and Isegaai T.  
 TITLE NESO human cDNA sequencing project  
 JOURNAL J Biol Chem  
 REFERENCE  
 AUTHORS Isegaai T. and Yamamoto J.  
 JOURNAL Direct Submission  
 TITLE Submitted (04-JUL-2002) Takao Isegaai, FIJ Project (HRI team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
 COMMENT (E-mail: genomics@ri.ken.go.jp; Tel:81-436-52-1975; Fax:81-436-52-3986) NESO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology (RAB); cDNA library construction; Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing; HRI and RAB; annotation: HRI and RAB.

FEATURES  
 Location/Qualifiers  
 1..3634  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="NT2RP7002594"  
 /cell\_line="H2"  
 /cell\_type="teratocarcinoma"  
 /clone\_id="H2H47"  
 /note="Cloning vector: pME18P3 mRNA from NT2 neuronal precursor cells after 5-weeks retinoic acid (RA) induction."

BASE COUNT 990 a 1014 c 805 g 825 t  
 ORIGIN  
 Query Match 100.0%; Score 216; DB 9; Length 3634;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-42;  
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCGCGCGCGCGCGAGGGTGGAGTCCATCCAGTGCACAGCTCCAGCATCTGGC 60  
 DB 2943 CCGCGCGCGCGCGCGAGGGTGGAGTCCATCCAGTGCACAGCTCCAGCATCTGGC 3002  
 QY 61 CAGCGCGTGAATGCTGAGCGACAGCGCTCAAGCCACAACTCACTGACAAAGTGG 120  
 DB 3003 CAGCGCGTGAATGCTGAGCGACAGCGCTCAAGCCACAACTCACTGACAAAGTGG 3062  
 QY 121 GGGCTGAAGCGTACGGCTGCTAAAGCGGAGCTACCGCCCAAGCACTCTTGGTCCG 180  
 DB 3063 GGGCTGAAGCGTACGGCTGCTAAAGCGGAGCTACCGCCCAAGCACTCTTGGTCCG 3122  
 QY 181 GTTCCATCATCATGAAGCCCAATGATGGGTGACA 216  
 DB 3123 GTTCCATCATCATGAAGCCCAATGATGGGTGACA 3158

RESULT 15  
 AX026746  
 LOCUS Sequence A from Parent, W00031252.  
 ACCESSION AX026746  
 VERSION 1  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Klosterman A. and Behl C.  
 TITLE Human semaphorin 6A-1 (sema6a-1), a gene involved in neuronal development and regeneration mechanisms during apoptosis, and its use as a potential drug target  
 JOURNAL Patent: WO 0031252-A 6, 02-JUN-2000;  
 Klosterman ANDREAS (DE) : MAX PLANCK GESELLSCHAFT (DE) : BEHL CHRISTIAN (DE)

FEATURES  
 Location/Qualifiers  
 1..3862  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 658..3750  
 /note="unlabeled protein product"  
 /protein\_id="CAC09109.1"  
 /db\_xref="GI:10187891"  
 /translation="MRSEALLPTLLRFAGAGPEDESEPTISISGNTKQYVPFVGH  
 KPGNTTORHLDIQMTIMNGTLYIAARDHIVTDIDTSHEIYCKSLTWKSRQA  
 DVDCRMKGKHKDECHNFIKVLKKDDALFVCGTNAFNPSCRYKMDLTPGDFES  
 CMACPYDAKXANVALFADGKLYSATVTDLAIDAVIYRISGESPLRTYKHDSKRIK  
 EYFVQAVDYGDIYFFREIAYNTMGVFPFPAQVCKNMGGQSQRVLEKQMTSF  
 LKALNCVPGDSHFYNILOAVTIVIRINGRDVLAFTSTPYNSIPGSACVAYMDL  
 -ASVTFGRFEKQSPDSTWTPVDPKPPGCCAGSSSLRYATISNEPDDTLNF  
 KTHPLMDEAVPSIFNRFWFLTMVRYRLTKIADTAAGPYQNHVTFVLSGEKGIILKF  
 LARINGSFLNDSLFLFLEMSVYSEKSYDQVDEKRMGMOLDRASSSLVAFSTCVI  
 KVLGKCEHCKKCTCIASDPYCGWKKEGGACSHLSPNSRLTFKQDIERGNTDGLG  
 DCHNSFVAINGHSSSLPSTTSDTAQEGESRGGLMDKHLHLDSPDPLCAVSS  
 INHQDKCVIRESIKGHDDQLVPVLLAIVLAFVMGAVFGITVTCVCDHRKKQVA  
 VYQREKELTISRKSGMSVTKSLGFDQSDKPKPEALTPLMHNGKLTATPNAK  
 MLIKADQHLDLTALPTPESTPQQKRPGRSGRWRNQNLIINACTKMPMPGSPV  
 IPTDLTPASPSHIPVYVVLPTIQOQYHFVQPKMSEVAQMALEDOAAATLYKRIK  
 EHLKSPNHCNVLNVDLSPKPKVPOREASLGPAGSLQGLSKHLEHHSSXGV  
 DYKRSYPTNSLTSRQHTTLKRNTNSNSHLSRNSQSGRGNPPAPORVDSIQVH  
 SQPSQGAQVTSRQPSLNAYNSLRSLKRPSPKDPVPPKPSFAPLSTSMKPNDACT

BASE COUNT 971 a 1111 c 967 g 813 t  
 ORIGIN

Query Match 100.0%; Score 216; DB 6; Length 3862;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-42;  
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

















GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CoreGen Inc.

OM protein - protein search, using sw model

Run on: September 10, 2003, 16:32:13 : Search time 445 seconds  
(without alignments)  
24.461 Million cell updates/sec

Title: US-09-856-681-4

Perfect score: 376

Sequence: 1 PPPAPQVDSIQVSSQPSG.....PTKSPAPISIMKPNDAIC 72

Scoring table: B:OSCM62

Gapop 10.0 , Gapext 0.5

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 566894

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

1: /cqn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cqn2\_6/prodata/2/pubpaa/PCR\_NEW\_PUB.pep.\*  
3: /cqn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cqn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cqn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
6: /cqn2\_6/prodata/2/pubpaa/PCRUS\_PUBCOMB.pep.\*  
7: /cqn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cqn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cqn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cqn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cqn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cqn2\_6/prodata/2/pubpaa/US09D\_PUBCOMB.pep.\*  
13: /cqn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cqn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cqn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cqn2\_6/prodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cqn2\_6/prodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cqn2\_6/prodata/2/pubpaa/US10F\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Score	Match	Length	DB ID	Description
1	376	100.0	1047	11	US-09-957-187-85 Sequence 85, Appl
2	86	22.9	873	12	US-09-864-761-34248 Sequence 18, Appl
3	80.5	21.4	374	12	US-10-239-431A-20 Sequence 20, Appl
4	80.5	21.4	594	12	US-10-239-431A-20 Sequence 8, Appl
5	74	19.7	629	15	US-10-213-956-36 Sequence 36, Appl
6	74	19.7	884	15	US-10-303-683-22 Sequence 21, Appl
7	74	19.7	884	15	US-09-291-417-18 Sequence 18, Appl
8	74	19.7	884	15	US-10-303-683-20 Sequence 20, Appl
9	73.5	19.5	2665	9	US-09-864-761-34248 Sequence 34248, A
10	73.5	19.5	3564	15	US-10-177-293-424 Sequence 423, Appl
11	73	19.4	200	12	US-10-259-165-300 Sequence 300, Appl
12	73	19.4	4019	10	US-09-738-973-425 Sequence 423, Appl
13	73	19.4	4019	10	US-09-854-133-425 Sequence 423, Appl
14	73	19.4	4019	15	US-10-144-649A-425 Sequence 423, Appl
15	71.5	19.0	223	9	US-09-864-761-33417 Sequence 33417, A

16 71.5 19.0 360 9 US-09-925-301-1116 Sequence 1116, Ap  
17 71.5 19.0 589 9 US-09-972-086-2 Sequence 2, Appl  
18 69.5 18.5 1125 15 US-10-100-957A-152 Sequence 152, App  
19 69.5 18.5 610 15 US-10-100-957A-22 Sequence 22, Appl  
20 68.5 18.2 625 10 US-09-801-368-294 Sequence 294, App  
21 68.5 18.2 729 15 US-10-145-396-11 Sequence 11, Appl  
22 68 18.1 153 10 US-09-925-300-1631 Sequence 1631, Ap  
23 68 18.1 632 12 US-09-934-455-122 Sequence 122, App  
24 68 18.1 697 15 US-10-218-957-2 Sequence 4, Appl  
25 68 18.1 612 15 US-10-218-957-2 Sequence 2, Appl  
26 68 18.1 1149 10 US-09-969-528-5 Sequence 5, Appl  
27 68 18.1 258 10 US-09-922-543-1 Sequence 1, Appl  
28 67.5 18.0 274 15 US-10-188-702A-6 Sequence 6, Appl  
29 67.5 18.0 465 15 US-10-153-668-218 Sequence 218, App  
30 67.5 18.0 5179 9 US-09-922-217-1069 Sequence 1068, Ap  
31 67.5 18.0 5179 10 US-09-833-253-1068 Sequence 1068, Ap  
32 67.5 18.0 5179 14 US-10-025-360-1068 Sequence 1068, Ap  
33 66.5 17.7 503 14 US-10-078-547-2 Sequence 2, Appl  
34 66.5 17.7 507 14 US-10-078-547-24 Sequence 24, Appl  
35 66 17.6 200 14 US-10-062-254-170 Sequence 170, App  
36 66 17.6 448 9 US-09-864-761-44230 Sequence 44230, A  
37 66 17.6 479 9 US-09-925-302-527 Sequence 527, App  
38 65.5 17.4 144 12 US-10-263-828-77 Sequence 77, Appl  
39 65.5 17.4 187 14 US-10-062-254-168 Sequence 168, App  
40 65.5 17.4 261 14 US-10-062-254-140 Sequence 140, App  
41 65.5 17.4 261 14 US-10-062-254-142 Sequence 142, App  
42 65.5 17.4 276 9 US-09-864-761-38306 Sequence 38306, A  
43 65.5 17.4 463 14 US-10-039-180-80 Sequence 80, Appl  
44 65.5 17.4 473 9 US-09-864-761-38321 Sequence 38321, A  
45 65.5 17.4 1142 10 US-09-899-651-2 Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-09-957-187-85  
: Sequence 85, Application US/C9957187  
: Publication No. US2003005414A1  
: GENERAL INFORMATION:  
: APPLICANT: Shimkets, Richard A.  
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY  
: FILE REFERENCE: 15966-540 CIP  
: CURRENT APPLICATION NUMBER: US/09/957,187  
: CURRENT FILING DATE: 2000-09-19  
: PRIOR APPLICATION NUMBER: 60/123,667  
: PRIOR FILING DATE: 1999-04-09  
: PRIOR APPLICATION NUMBER: 09/520,761  
: PRIOR FILING DATE: 2000-03-03  
: PRIOR APPLICATION NUMBER: 60/234,082  
: PRIOR FILING DATE: 2000-09-20  
: PRIOR APPLICATION NUMBER: 60/233,798  
: PRIOR FILING DATE: 2000-09-19  
: PRIOR APPLICATION NUMBER: 60/174,485  
: PRIOR FILING DATE: 2000-01-04  
: NUMBER OF SEQ ID NOS: 85  
: SOFTWARE: Patentin Ver. 2.1  
: SEQ ID NO 85  
: LENGTH: 1047  
: TYPE: PRT  
: ORGANISM: Homo sapiens  
US-09-957-187-85

Query Match 100.0% Score 376; DB 11; Length 1047;  
Best Local Similarity 100.0% Pred. No. 1.9e-29;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQVDSIQVSSQPSGQAVTVSRQPSLNAYNSLITRSLKRTPLSKPVPKPSFAP 60  
DB 976 PPPAPQVDSIQVSSQPSGQAVTVSRQPSLNAYNSLITRSLKRTPLSKPVPKPSFAP 1035  
CY 61 LSTSMKPNDAICT 72

Db 1036 LSTSMKP 67  
 259 VAPKPT 265

## RESULT 2

US-09-843-245-3  
 : Sequence 3, Application US/09843245  
 : Patent No. US20020164672A1

## GENERAL INFORMATION:

: APPLICANT: McPherson, Peter S.  
 : APPLICANT: Ramjaun, Antoine Rachid  
 : TITLE OF INVENTION: REGULATION OF JNK ACTIVITY BY MODULATION OF THE  
 : INTERACTION BETWEEN THE ENDOCYTIC PROTEIN ENDOPHILIN  
 : AND THE GERMINAL CENTER KINASE-LIKE KINASE  
 : FILE REFERENCE: 9555.1166501  
 : CURRENT APPLICATION NUMBER: US/09/843,245  
 : CURRENT FILING DATE: 2001-04-25

: NUMBER OF SEQ ID NOS: 13

: SOFTWARE: Patent In Ver. 2.1

: SEQ ID NO 3

: LENGTH: 873

: TYPE: PRT

: ORGANISM: Rattus sp.

US-09-843-245-3

Query Match 22.9%; Score 86; DB 10; Length 873;

Best Local Similarity 34.7%; Pred. No. 2;

Matches 25; Conservative 6; Mismatches 22; Indels 14; Gaps 3;

QY 1 PPAPQVDSGVHSSQPSQAVTVSRPSLKNAYNSLTRSGKRTSLKPVPPKPSFAP 56

DB 41: PPQVQAKPTQVQVH-VQQAQP-HVQQP-VSSANTQPKPLSQAPT-----PAPKPT 465

QY 57 SEAPLSTSMKN 68

DB 44: RLPPQKAVLGN 472

## RESULT 3

US-10-239-431A-20  
 : Sequence 20, Application US/10139431A  
 : Publication No. US20030170725A1

## GENERAL INFORMATION:

: APPLICANT: FRADELIZE, JULIE  
 : APPLICANT: FRIEDERICH, EVELYNE  
 : APPLICANT: GOLSTEYN, ROY M.  
 : APPLICANT: LOUWARD, DANIEL  
 : APPLICANT: NOIREAUX, VINCENT  
 : APPLICANT: SYKES, CECILE

: TITLE OF INVENTION: PEPTIDE SEQUENCES COMPRISING ONE OR MORE UNITS BINDING

: TO PROTEINS OF THE ENA/VASP FAMILY, AND THEIR USES

: FILE REFERENCE: 0508-1032

: CURRENT APPLICATION NUMBER: US/10/239,431A

: PRIOR FILING DATE: 2002-09-23

: PRIOR APPLICATION NUMBER: PCT/FR01/00843

: PRIOR FILING DATE: 2001-03-21

: PRIOR APPLICATION NUMBER: FR 00/03637

: NUMBER OF SEQ ID NOS: 38

: SOFTWARE: Patent In Ver. 2.1

: SEQ ID NO 20

: LENGTH: 374

: TYPE: PRT

: ORGANISM: Murine sp.

US-10-239-431A-20

Query Match 21.4%; Score 80.5; DB 12; Length 374;

Best Local Similarity 32.8%; Pred. No. 2,6;

Matches 22; Conservative 11; Mismatches 25; Indels 9; Gaps 4;

QY 1 PPAPQVDSGVHSSQPSQAVTVSRPSLKNAYNSLTRSGKRTSLKPVPPKPSFAP 60

DB 208 PPQVQAKPTQVQVH-VQQAQP-HVQQP-VSSANTQPKPLSQAPT-----PAPKPT 258

QY 61 LSTSMKP 67  
 259 VAPKPT 265

## RESULT 4

US-10-239-431A-8  
 : Sequence 8, Application US/10239431A  
 : Publication No. US20030170726A1

## GENERAL INFORMATION:

: APPLICANT: FRADELIZE, JULIE  
 : APPLICANT: FRIEDERICH, EVELYNE  
 : APPLICANT: GOLSTEYN, ROY M.  
 : APPLICANT: LOUWARD, DANIEL  
 : APPLICANT: NOIREAUX, VINCENT  
 : APPLICANT: SYKES, CECILE

: TITLE OF INVENTION: PEPTIDE SEQUENCES COMPRISING ONE OR MORE UNITS BINDING

: TO PROTEINS OF THE ENA/VASP FAMILY, AND THEIR USES

: FILE REFERENCE: 0508-1032

: CURRENT APPLICATION NUMBER: US/10/239,431A

: CURRENT FILING DATE: 2002-09-23

: PRIOR APPLICATION NUMBER: PCT/FR01/00843

: PRIOR FILING DATE: 2001-03-21

: PRIOR APPLICATION NUMBER: FR 00/03637

: PRIOR FILING DATE: 2000-03-22

: NUMBER OF SEQ ID NOS: 38

: SOFTWARE: Patent In Ver. 2.1

: SEQ ID NO 8

: LENGTH: 564

: TYPE: PRT

: ORGANISM: Murine sp.

US-10-239-431A-8

Query Match 21.4%; Score 80.5; DB 12; Length 564;

Best Local Similarity 32.8%; Pred. No. 4,2;

Matches 22; Conservative 11; Mismatches 25; Indels 9; Gaps 4;

QY 1 PPAPQVDSGVHSSQPSQAVTVSRPSLKNAYNSLTRSGKRTSLKPVPPKPSFAP 60

DB 209 PPQVQAKPTQVQVH-VQQAQP-HVQQP-VSSANTQPKPLSQAPT-----PAPKPT 259

QY 61 LSTSMKP 67

DB 260 VAPKPT 265

## RESULT 5

US-10-213-990-36

: Sequence 36, Application US/10213990

: Publication No. US20030082595A1

## GENERAL INFORMATION:

: APPLICANT: Jiang, Bo

: APPLICANT: Bussey, Howard

: APPLICANT: Storms, Terry

: APPLICANT: Roemer, Reg

: TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL

: ENZYMS AND METHODS OF USE

: FILE REFERENCE: 10182-019-999

: CURRENT APPLICATION NUMBER: US/10/213,990

: CURRENT FILING DATE: 2002-08-05

: NUMBER OF SEQ ID NOS: 72

: SOFTWARE: FastSeq for Windows version 4.0

: SEQ ID NO 36

: LENGTH: 628

: TYPE: PRT

: ORGANISM: Aspergillus

US-10-213-990-36

Query Match 19.7%; Score 74; DB 15; Length 628;

Best Local Similarity 42.2%; Pred. No. 22;

Matches 19; Conservative 5; Mismatches 19; Indels 2; Gaps 1;

```
QY 6 QVDSQVSSQSGAVTVSRQPSLNAYNSL--TRSGLKRTPS 48
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 178 QVDSNPINFGAPAGLEVTRDPTVNCWPSLRAGQGVASDQ 222
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-10-303-683-21
; Sequence 21, Application US/10-04583
; Publication No. US20030113782A1
; GENERAL INFORMATION:
; APPLICANT: Exelixis, Inc.
; TITLE OF INVENTION: MAP4KS AS MODIFIERS OF BRANCHING MORPHOGENESIS AND METHODS OF USE
; FILE REFERENCE: EXG2-126C
; CURRENT APPLICATION NUMBER: US/10/303,583
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 60/333,378
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 884
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-303-683-21

Query Match 19.7%; Score 74; DB 15; Length 884;
Best Local Similarity 32.8%; Pred. No. 32;
Matches 21; Conservative 6; Mismatches 23; Indels 14; Gaps 3;

QY 1 PPPAFQVDSI---QVHSQPSGQAVTVSRQPSLNAYNSLRSGLKRTPSLKPDVPPKP 56
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 PPLPKPKSIPTQPMHSTEDENOG-TTKRCP-----MSGSPAKPSQVPPRPPPP 473
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 57 SFAP 60
Db 472 RLPP 475

RESULT 7
US-09-291-417-18
; Sequence 18, Application US/09291417A
; Publication No. US2003050230A1
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: MARINEZ, RICARDO
; APPLICANT: WEXEL, DAVID
; TITLE OF INVENTION: STR20-RELATED PROTEIN KINASES
; FILE REFERENCE: 249/300
; CURRENT APPLICATION NUMBER: US/09/291,417A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: US 60/081,784
; EARLIER FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Mammalian (Human) KHS2
US-09-291-417-18

Query Match 19.7%; Score 74; DB 15; Length 894;
Best Local Similarity 32.8%; Pred. No. 32;
Matches 21; Conservative 6; Mismatches 23; Indels 14; Gaps 3;

QY 1 PPPAFQVDSI---QVHSQPSGQAVTVSRQPSLNAYNSLRSGLKRTPSLKPDVPPKP 56
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 432 PPLPKPKSIPTQPMHSTEDENOG-TTKRCP-----MSGSPAKPSQVPPRPPPP 481
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 57 SFAP 60
Db 482 RLPP 485

RESULT 8
US-10-303-683-20
; Sequence 20, Application US/10303683
; Publication No. US20030113782A1
; GENERAL INFORMATION:
; APPLICANT: Exelixis, Inc.
; TITLE OF INVENTION: MAP4KS AS MODIFIERS OF BRANCHING MORPHOGENESIS AND METHODS OF USE
; FILE REFERENCE: EXG2-126C
; CURRENT APPLICATION NUMBER: US/10/303,683
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 60/333,378
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-303-683-20

Query Match 19.7%; Score 74; DB 15; Length 894;
Best Local Similarity 32.8%; Pred. No. 32;
Matches 21; Conservative 6; Mismatches 23; Indels 14; Gaps 3;

QY 1 PPPAFQVDSI---QVHSQPSGQAVTVSRQPSLNAYNSLRSGLKRTPSLKPDVPPKP 56
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 432 PPLPKPKSIPTQPMHSTEDENOG-TTKRCP-----MSGSPAKPSQVPPRPPPP 481
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 57 SFAP 60
Db 482 RLPP 485

RESULT 9
US-09-864-761-34248
; Sequence 34248, Application US/09864761
; Patent No. US20020348763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Shartor G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
```



Best Local Similarity 30.4%; Pred. No. 7.3;  
Matches 21; Conservative 11; Mismatches 33; Indels 4; Gaps 3;

QY 2 PPAPQVVSQVES-SQSSQAVTVSROPSLNAYNSLTRSLKRTTP-----45  
DB 129 PAVQAVQVQCSAISSPPAFISTCKEAPRVYVETKATRRKSPATTFSPAPPASR 186  
QY 58 EAPLSTSMK 66  
DB 189 AAFCASSER 157

## RESULT 12

US-09-738-973-425  
Sequence 425, Application US/0373553  
Patent No. US2002010563A1  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Henderson, Robert A.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Fling, Steven P.  
APPLICANT: Monamath, Raouleh  
APPLICANT: Algate, Paul A.  
APPLICANT: Secrist, Heather  
APPLICANT: Indrias, Carol Yoseph  
APPLICANT: Benson, Darin R.  
APPLICANT: Elliot, Mark  
APPLICANT: Mannion, Jane  
APPLICANT: Kalos, Michael D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.475C9  
CURRENT APPLICATION NUMBER: US/09/738.973  
CURRENT FILING DATE: 2000-12-14  
NUMBER OF SEQ ID NOS: 587  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 425  
LENGTH: 4019  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-738-973-425

Query Match 19.4%; Score 73; DB 16; Length 4019;  
Best Local Similarity 32.5%; Pred. No. 2.3e+02;  
Matches 27; Conservative 11; Mismatches 27; Indels 14; Gaps 6;

QY 1 PPAPQVVSQVES-SQSSQAVTVSROPSLNAYNSLTRSLKRTTP-----45  
DB 910 PPAPSRIFQDSLSQACTSQSPQVSPGSSNSRPPSPDPYAKMV--GTPRPDPVGH 967  
QY 47 --SLKPDVPPKPSFAPISTSMKP 67  
DB 968 SFSRNSAAPVENCITLSSVSRP 990

## RESULT 13

US-09-854-133-425  
Sequence 425, Application US/09854133  
Patent No. US20020183499A1  
GENERAL INFORMATION:  
APPLICANT: Lodes, Michael J.  
APPLICANT: Monamath, Raouleh  
APPLICANT: Henderson, Robert A.  
APPLICANT: Benson, Darin R.  
APPLICANT: Secrist, Heather  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.475C1C  
CURRENT APPLICATION NUMBER: US/09/854.133  
CURRENT FILING DATE: 2001-05-11  
NUMBER OF SEQ ID NOS: 735  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 425

LENGTH: 4019  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-854-133-425

Query Match 19.4%; Score 73; DB 10; Length 4019;  
Best Local Similarity 32.5%; Pred. No. 2.3e+02;  
Matches 27; Conservative 11; Mismatches 27; Indels 18; Gaps 6;

QY 1 PPAPQVVSQVES-SQSSQAVTVSROPSLNAYNSLTRSLKRTTP-----45  
DB 910 PPAPSRIFQDSLSQACTSQSPQVSPGSSNSRPPSPDPYAKMV--GTPRPDPVGH 967  
QY 47 --SLKPDVPPKPSFAPISTSMKP 67  
DB 968 SFSRNSAAPVENCITLSSVSRP 990

## RESULT 14

US-10-144-649A-425  
Sequence 425, Application US/0144649A  
Publication No. US2003011859A1  
GENERAL INFORMATION:  
APPLICANT: Lodes, Michael J.  
APPLICANT: Wang, Tongtong  
APPLICANT: Fan, Liqun  
APPLICANT: Algate, Paul A.  
APPLICANT: McNeill, Patricia D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.475C11  
CURRENT APPLICATION NUMBER: US/10/-44.649A  
CURRENT FILING DATE: 2002-08-21  
NUMBER OF SEQ ID NOS: 749  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 425  
LENGTH: 4019  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-144-649A-425

Query Match 19.4%; Score 73; DB 15; Length 4019;  
Best Local Similarity 32.5%; Pred. No. 2.3e+02;  
Matches 27; Conservative 11; Mismatches 27; Indels 18; Gaps 6;

QY 1 PPAPQVVSQVES-SQSSQAVTVSROPSLNAYNSLTRSLKRTTP-----46  
DB 910 PPAPSRIFQDSLSQACTSQSPQVSPGSSNSRPPSPDPYAKMV--GTPRPDPVGH 967  
QY 47 --SLKPDVPPKPSFAPISTSMKP 67  
DB 968 SFSRNSAAPVENCITLSSVSRP 990

## RESULT 15

US-09-864-751-33417  
Sequence 33417, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aecmica-X-1  
CURRENT APPLICATION NUMBER: US/09/864.761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366

```

1 PRIOR FILING DATE: 2000-08-03
2 PRIOR APPLICATION NUMBER: GB 24263.6
3 PRIOR FILING DATE: 2000-10-04
4 PRIOR APPLICATION NUMBER: US 60/236,359
5 PRIOR FILING DATE: 2000-09-27
6 PRIOR APPLICATION NUMBER: PCT/US01/00666
7 PRIOR FILING DATE: 2001-01-30
8 PRIOR APPLICATION NUMBER: PCT/US01/00667
9 PRIOR FILING DATE: 2001-01-30
10 PRIOR APPLICATION NUMBER: PCT/US01/00664
11 PRIOR FILING DATE: 2001-01-30
12 PRIOR APPLICATION NUMBER: PCT/US01/00669
13 PRIOR FILING DATE: 2001-01-30
14 PRIOR APPLICATION NUMBER: PCT/US01/00665
15 PRIOR FILING DATE: 2001-01-30
16 PRIOR APPLICATION NUMBER: PCT/US01/00668
17 PRIOR FILING DATE: 2001-01-30
18 PRIOR APPLICATION NUMBER: PCT/US01/00663
19 PRIOR FILING DATE: 2001-01-30
20 PRIOR APPLICATION NUMBER: PCT/US01/00662
21 PRIOR FILING DATE: 2001-01-30
22 PRIOR APPLICATION NUMBER: PCT/US01/00661
23 PRIOR FILING DATE: 2001-01-30
24 PRIOR APPLICATION NUMBER: PCT/US01/00660
25 PRIOR FILING DATE: 2001-01-30
26 PRIOR APPLICATION NUMBER: US 60/234,687
27 PRIOR FILING DATE: 2000-09-27
28 PRIOR APPLICATION NUMBER: US 09/608,438
29 PRIOR FILING DATE: 2000-06-30
30 PRIOR APPLICATION NUMBER: US 09/774,204
31 PRIOR FILING DATE: 2001-01-29
32 NUMBER OF SEQ ID NOS: 49117
33 SOFTWARE: Ancomax Sequence Listing Engine vers. 1.1
34 SEQ ID NO 33417
35 LENGTH: 223
36 TYPE: PRI
37 ORGANISM: Homo sapiens
38 FEATURE:
39 OTHER INFORMATION: MAP TO AC006581.16
40 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 13
41 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 41
42 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 25
43 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 25
44 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 28
45 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 46
46 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 17
47 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 71
48 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 15
49 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 29
50 OTHER INFORMATION: EST_HUMAN HT: AW959289.1, EVALU 9.00e-15
51 US-09-864-761-33417

```

```

Query Match 19.0% Score 71.5; DB 9; Length 223;
Best Local Similarity 35.8%; Pred No. 12;
Matches 24; Conservative 3; Mismatches 23; Indels 15; Gaps 3;

QY 2 PPAPORVSIQVHSCPGQAVTVSRQPSLNAYNS:ATSG:KRTF-SLKPVPEKPS:AP 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 79 PGPQVPTQCVPPSQSQQACTLVVQPMLOS-----SPLSLPDAAPKP---P 124
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 LSTSMKP 67
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 125 IPIQSKP 131

```

Search completed: September 30, 2003, 16:43:58  
Job time : 449 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search, using sw mode.

Run on: September 30, 2003, 13:15:08 : Search time 72 seconds  
(without alignments)  
158,747 million cells updated/sec

Title: US-09-856-681-4

Perfect score: 376

Sequence: 1 PPPAFQVDSICVHSSQPSG.....PPQSFAPLSTSMKPLNDACT 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 6.5

Searched: 1107663 seqs, 159726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	DB ID	Description
1	376	100.0	72	21: AAV71461 Binding domain of
2	376	100.0	507	22: AAB94288 Human protein seq
3	376	100.0	562	22: AAM93444 Human polypeptide
4	376	100.0	562	22: AAB94104 Human protein seq
5	376	100.0	574	22: AAB95317 Human protein seq
6	376	100.0	599	22: AAB95139 Human protein seq
7	376	100.0	863	22: ASG04066 Novel human diag
8	376	100.0	974	19: AAW64221 Human secreted pro
9	376	100.0	975	22: AAB90731 Human G3145_1 pro

10	376	100.0	1050	21: AAV71460 Human semaphorin 6
11	163.5	43.5	451	22: AAB94239 Human protein seq
12	163.5	43.5	464	22: AAB94286 Human protein seq
13	163.5	43.5	474	24: ABU11724 Human MBDT polyep
14	163.5	43.5	1017	23: AAG79413 CAPUP-2, InCyte IL
15	163.5	43.5	1032	23: AB379125 Human semaphorin-1
16	163.5	43.5	1035	23: AB379173 Human semaphorin 6
17	163.5	43.5	1068	22: AAG63213 Arino acid sequenc
18	163.5	43.5	1082	23: ABG79177 Human semaphorin-1
19	86	22.4	473	24: AP174448 Rat germinall centr
20	80.5	21.4	374	22: AAM52309 Murine zyxine frag
21	80.5	21.4	564	22: AAM52303 Murine zyxine, M
22	80.5	21.4	564	23: ABB57260 Mouse ischaemic co
23	77.5	20.6	693	23: ABP95529 Human polypeptide
24	77.5	20.6	1111	23: ABB93287 Herbicidally activ
25	75.5	20.1	735	22: ABB65233 Prosophila melanog
26	75.5	20.1	745	22: ABB65321 Prosophila melanog
27	75	19.9	95	22: AAU45062 Propionibacterium
28	74.5	19.8	449	22: ABB60133 Prosophila melanog
29	74.5	19.8	1475	22: ABB71432 Prosophila melanog
30	74	19.7	873	22: AAM78346 Human protein SEQ
31	74	19.7	894	22: AAY55935 Human KHS2 protein
32	74	19.7	894	22: AAM78345 Human protein SEQ
33	74	19.7	930	22: AAM79329 Human protein SEQ
34	74	19.7	930	22: AAM79330 Human protein SEQ
35	73.5	19.5	425	22: ABB22456 Novel human diagno
36	73.5	19.5	960	22: ASG20323 Novel human diagno
37	73.5	19.5	960	22: ABG20886 Novel human diagno
38	73.5	19.5	2665	22: ASG48336 Human liver peptid
39	73.5	19.5	2665	22: ASB28314 Human peptide #965
40	73.5	19.5	2665	22: ABB33450 Peptide #996 encod
41	73.5	19.5	2665	22: ABB18950 Protein #949 encod
42	73.5	19.5	2665	22: AAM54270 Human brain expres
43	73.5	19.5	2665	22: AAM66665 Human bone marrow
44	73.5	19.5	2665	22: AAM14533 Peptide #967 encod
45	73.5	19.5	2665	22: AAM26950 Peptide #987 encod

#### ALIGNMENTS

#### RESULT 1

AAV71461

10 AAV71461 standard; peptide: 72 AA.

XX AAV71461;

XX 04-OCT-2003 (first entry)

XX Binding domain of human semaphorin 6A-1.

XX Human: semaphorin 6A-1; (HSA)SEMA6A-1; neuronal development; apoptosis;  
neural regeneration; Fna/VASP protein family; immunomodulatory;  
neurotrophic agent; therapeutic agent; differentiation;  
cytoskeletal stabilisation; plasticity.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Binding-site 51..56

XX FT /note= "Specific binding motif for members of  
FT Fna/VASP protein family, especially Evi"

XX WO200031252-A1.

XX 02-JUN-2003.

XX 26-NOV-1993; 99WO-EP0922.5.

XX 26-NOV-1998; 98EP-0122441.

XX (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PI Behl C, Klostermann A;  
 XX WPI: 2000-400065/34.  
 DR N-PSDB: AAD01234.  
 XX  
 PT Nucleic acid coding for human semaphorin 6A-1 used as diagnostic agent,  
 PT therapeutic agent, for modulating immune system, in gene therapy or for  
 PT effecting differentiation, cytoskeletal stabilization and/or plasticity  
 PT -  
 XX  
 PS Disclosure: Page 22; 53pp; English.  
 XX  
 CC The present sequence is a binding domain of transmembranes  
 CC human semaphorin 6A-1 ((HSA)SEMA6A-1) which is involved in neuronal  
 CC development and regeneration mechanisms during apoptosis. The binding  
 CC domain shows homology to zyxin protein and selectively binds to members  
 CC of Ena/VASP protein family, especially Evi1. (HSA)SEMA6A-1 is a  
 CC member of protein family displaying secreted or transmembrane-based  
 CC repulsive guidance cues critically involved in neuronal development.  
 CC Expression of (HSA)SEMA6A-1 is highest in embryonic brain and  
 CC kidney and moderate in lung. The present sequence is useful as diagnostic  
 CC and therapeutic agents for modulating the immune system, in gene  
 CC therapy, for effecting differentiation, cytoskeletal stabilisation  
 CC and plasticity.  
 XX  
 SQ Sequence 72 AA:  
 Query Match 100.0%; Score 376; DB 21; Length 72;  
 Best Local Similarity 100.0%; Pred. No. 7.4e-25;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPPAPQVDSIQVHSSQSGQAVTVSRGSLKTPSLKTPVPPKPSFAP 60  
 DB 1 PPPAPQVDSIQVHSSQSGQAVTVSRGSLKTPSLKTPVPPKPSFAP 60  
 QY 61 LSTSMKPNDACT 72  
 DB 61 LSTSMKPNDACT 72  
 RESULT 2  
 AAB92688  
 ID AAB92688 standard; Protein: 507 AA.  
 XX  
 AC AAB92688;  
 XX  
 DT 26-JUN-2001 (first entry)  
 DE Human protein sequence SEQ ID NO:11073.  
 XX  
 DE Human; primer: detection; diagnosis; antisense therapy; gene therapy.  
 KW Homo sapiens.  
 OS  
 PN EP1074417-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito X, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI: 2001-318749/34.

PI Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 CC Claim 8; SEQ ID 11073; 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC particularly full-length cDNAs. The primers are useful for synthesizing  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAB03166 to AAB13628 and  
 CC AAB13633 to AAB18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAB13629 to AAB13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SQ Sequence 507 AA:  
 Query Match 100.0%; Score 376; DB 22; Length 507;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-34;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPPAPQVDSIQVHSSQSGQAVTVSRGSLKTPSLKTPVPPKPSFAP 60  
 DB 496 PPPAPQVDSIQVHSSQSGQAVTVSRGSLKTPSLKTPVPPKPSFAP 495  
 QY 61 LSTSMKPNDACT 72  
 DB 496 LSTSMKPNDACT 507  
 RESULT 3  
 AAB93444  
 ID AAB93444 standard; Protein: 562 AA.  
 XX  
 AC AAB93444;  
 XX  
 DT 06-NOV-2001 (first entry)  
 DE Human polypeptide. SEQ ID NO: 3088.  
 XX  
 DE Human; full length cDNA; cDNA synthesis; oligo-capping.  
 KW Homo sapiens.  
 OS  
 PN EP110094-A2.  
 XX  
 PD 05-SEP-2001.  
 XX  
 PF 07-JUL-2000; 2000EP-0114089.  
 XX  
 PR 09-JUL-1999; 99JP-0194486.  
 PR 11-JAN-2000; 2000JP-0118774.  
 PR 02-MAY-2000; 2000JP-0183765.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koqa H;



XX WPI: 2001-524355/58.  
 DR R-PSDS: AAK94365.  
 XX  
 PT 830 Primers useful for synthesizing full length cDNA clones and their  
 use in genetic manipulation.  
 XX  
 PS Claim 8: SEQ ID NO 3088: 1380pp + sequence listing: English.  
 XX  
 CC The invention relates to primers for synthesizing full length cDNA  
 clones. 830 cDNA molecules encoding a human protein have been  
 isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
 molecules have been determined. Primers for synthesizing the full length  
 cDNA are useful for clarifying the function of the protein encoded by  
 the cDNA. The full length clones were obtained by construction of full  
 length enriched cDNA libraries that were synthesised by the oligo-repairing  
 method. The primers enable the production of the full length cDNA easily  
 without any special methods. The present sequence is a polypeptide  
 encoded by a full length human cDNA of the invention.  
 CC Note: The sequence data for this patent did not form part of the claimed  
 specification, but was obtained in CD-ROM format directly from EPO.  
 XX  
 SQ Sequence 562 AA:  
 Query Match 100.0%; Score 376; DB 22; Length 562;  
 Best Local Similarity 100.0%; Pred. No. 9,4e-34;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60  
 DB 491 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 560  
 QY 61 LSTSMKPNDACT 72  
 DB 55 LSTSMKPNDACT 562  
 RESULT 4  
 AAB94104  
 ID AAB94104 standard; Protein: 562 AA.  
 XX AAB94104:  
 XX 26-JUN-2001 (first entry)  
 XX Human protein sequence SEQ ID NO:14328.  
 XX Human: primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX Homo sapiens.  
 XX EP1074617-A2.  
 XX 07-FEB-2001.  
 XX 28-JUL-2000; 2000EP-0116126.  
 XX 29-JUL-1999; 99JP-0248036.  
 XX 27-AUG-1999; 99JP-0300253.  
 XX 11-JAN-2000; 2000JP-0118776.  
 XX 02-MAY-2000; 2000JP-0183767.  
 XX 09-JUN-2000; 2000JP-0241699.  
 XX (HELI-) HELIX RES INST.  
 XX Ota T, Isozaki T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI: 2001-318749/34.  
 DR  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 full-length cDNAs defined in the specification, and for the detection  
 and/or diagnosis of the abnormality of the proteins encoded by the

PI full-length cDNAs -  
 XX Claim 8: SEQ ID 14328, 2537pp - CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602  
 full-length cDNAs defined in the specification. Where a primer set  
 comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 to the complementary strand of a polynucleotide which comprises one of  
 the 5602 nucleotide sequences defined in the specification, where the  
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 of an oligonucleotide comprising a sequence complementary to the  
 complementary strand of a polynucleotide which comprises a 5'-end  
 sequence and an oligonucleotide comprising a sequence complementary to a  
 polynucleotide which comprises a 3'-end sequence, where the  
 oligonucleotide comprises at least 15 nucleotides and the combination of  
 the 5'-end sequence/3'-end sequence is selected from those defined in  
 the specification. The primer sets can be used in antisense therapy and  
 in gene therapy. The primers are useful for synthesizing polynucleotides,  
 particularly full-length cDNAs. The primers are also useful for the  
 detection and/or diagnosis of the abnormality of the proteins encoded by  
 the full-length cDNAs. The primers allow obtaining of the full-length  
 cDNAs easily without any specialised methods. AAB93166 to AAB93628 and  
 AAB93633 to AAB93742 represent human cDNA sequences; AAB92446 to  
 AAB95893 represent human amino acid sequences; and AAB93629 to AAB93632  
 represent oligonucleotides, all of which are used in the exemplification  
 of the present invention.  
 XX  
 SQ Sequence 562 AA:  
 Query Match 100.0%; Score 376; DB 22; Length 562;  
 Best Local Similarity 100.0%; Pred. No. 9,4e-34;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60  
 DB 491 PPPAPQRVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 560  
 QY 61 LSTSMKPNDACT 72  
 DB 55 LSTSMKPNDACT 562  
 RESULT 5  
 AAB95317  
 ID AAB95317 standard; Protein: 574 AA.  
 XX AAB95317:  
 XX 26-JUN-2001 (first entry)  
 XX Human protein sequence SEQ ID NO:17566.  
 XX Human: primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX Homo sapiens.  
 XX EP1074617-A2.  
 XX 07-FEB-2001.  
 XX 28-JUL-2000; 2000EP-0116126.  
 XX 29-JUL-1999; 99JP-0248036.  
 XX 27-AUG-1999; 99JP-0300253.  
 XX 11-JAN-2000; 2000JP-0118776.  
 XX 02-MAY-2000; 2000JP-0183767.  
 XX 09-JUN-2000; 2000JP-0241699.  
 XX (HELI-) HELIX RES INST.  
 XX Ota T, Isozaki T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI: 2001-318749/34.

DR WPI: 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs.  
 XX  
 PS Claim 8: SEQ ID 17568: 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC polynucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primers are useful for synthesizing polynucleotides,  
 CC in gene therapy. The primers can be used in antisense therapy and  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SQ Sequence 574 AA:  
 Query Match 100.0%; Score 376; Db 22; Length 574;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-34;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPAPAPRVNISTGVHSSQPSGQAVTVSRQPSLNAYNSLTRSGIKRTPSLKPDVPPKPSFAP 60  
 DB 503 PPAPAPRVNISTGVHSSQPSGQAVTVSRQPSLNAYNSLTRSGIKRTPSLKPDVPPKPSFAP 582  
 QY 61 LSTSMKPNDACT 72  
 DB 563 LSTSMKPNDACT 574  
 RESULT 6  
 AAB95139  
 ID AAB95139 standard; Protein: 599 AA.  
 XX  
 AC AAB95139;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human protein sequence SEQ ID NO: 7154.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0249036.  
 PR 27-AUG-1999; 99JP-030253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0133767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX

PA (HELIX) HELIX RES INCT.  
 XX  
 PI Ora T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakatsuki A, Nagai K, Otsuki T;  
 XX  
 DB WPI: 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs.  
 XX  
 PS Claim 8: SEQ ID 17154: 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primers are useful for synthesizing polynucleotides,  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SQ Sequence 699 AA:  
 Query Match 100.0%; Score 376; Db 22; Length 699;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-33;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PPAPAPRVNISTGVHSSQPSGQAVTVSRQPSLNAYNSLTRSGIKRTPSLKPDVPPKPSFAP 60  
 DB 628 PPAPAPRVNISTGVHSSQPSGQAVTVSRQPSLNAYNSLTRSGIKRTPSLKPDVPPKPSFAP 587  
 QY 61 LSTSMKPNDACT 72  
 DB 686 LSTSMKPNDACT 699  
 RESULT 7  
 AAB94066  
 ID AAB94066 standard; Protein: 863 AA.  
 XX  
 AC AAB94066;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #4057.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US08631.  
 XX

```

PR 31-MAR-2000: 2500US-2540217.
PR 23-AUG-2000: 2500US-2619157.
XX (HYSE-) HYSEQ INC.
XX Dmanac RT, Liu C, Tang YD:
XX
XX
XX
XX WPI: 2001-639162/73.
DR N-PSDB: AAS68253.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
XX Claim 20: SEQ ID No 34425; 103pp: English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAG000-AB350377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pt_sequences.
XX
XX Sequence 863 AA:
XX
XX Query Match 100.0%; Score 376; DB 22; Length 863;
XX Best Local Similarity 100.0%; Pred. No. 1,56-33;
XX Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 PPPAPQVDSIQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
XX '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 792 PPPAPQVDSIQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 450
XX
XX QY 61 LSTSMKPNDACT 72
XX '|||||:|||||
XX DB #52 LSTSMKPNDACT 863
XX
XX RESULT 8
XX AAW64221
XX ID AAW64221 standard; Protein: 974 AA.
XX
XX AC AAW64221:
XX
XX DT 06-OCT-1998 (first entry)
XX
XX DE Human secreted protein from clone CJ145_1.
XX
XX KW Secreted protein; human fetal brain; nutrition; cytokine; stimulant;
XX cell proliferation; differentiation; immune system; suppressor; ligand;
XX regulator; hematopoiesis; tissue growth; activin; inhibin; haemostatic;
XX chemotaxis; chemokine; thrombosis; receptor; cadherin; tumour;
XX anti-inflammatory.
XX
XX OS Homo sapiens.
XX
XX PN WO9827205-A2.

```

---

```

XX 25-JUN-1998.
XX
XX 17-DEC-1997: 97W03-US23330.
XX
XX 16-DEC-1997: 97US-0591872.
XX 18-DEC-1996: 96US-0749392.
XX 14-JAN-1997: 97US-0783401.
XX
XX (GENY) GENETICS INST INC.
XX
XX Agostino M, Jacobs K, Lavallie ER, McCoy JM, Merberg D:
XX Racie LA, Spaulding V, Treacy M;
XX
XX WPI: 1998-362774/31.
XX N-PSDB: AAV44295.
XX
XX New polynucleotides and secreted proteins - obtained from human:
XX foetal brain, human adult testes, human adult brain and human adult
XX salivary gland cDNA libraries
XX
XX Claim 17j: Page 71-74; 110pp: English.
XX
XX This sequence represents a novel secreted protein from clone CJ145_1
XX isolated from a human total brain cDNA library. This protein has
XX applications for nutritional use, cytokine and cell
XX proliferation/differentiation activity, immune stimulating or
XX suppressing activity, hematopoiesis regulating activity, tissue growth
XX activity, activin/inhibin activity, chemotactic/chemokinetic activity,
XX haemostatic and thrombotic activity, receptor/ligand activity,
XX anti-inflammatory activity, cadherin/tumour invasion suppressor activity,
XX tumour inhibition activity and other activities.
XX
XX Sequence 974 AA:
XX
XX Query Match 100.0%; Score 376; DB 19; Length 974;
XX Best Local Similarity 100.0%; Pred. No. 1,9e-33;
XX Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 PPPAPQVDSIQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
XX '|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 903 PPPAPQVDSIQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 562
XX
XX QY 61 LSTSMKPNDACT 72
XX '|||||:|||||
XX DB 63 LSTSMKPNDACT 974
XX
XX RESULT 9
XX AAW90731
XX ID AAW90731 standard; Protein: 975 AA.
XX
XX AC AAW90731:
XX
XX DT 07-JUN-2001 (first entry)
XX
XX DE Human CJ145_1 protein sequence SEQ ID 161.
XX
XX KW Human; secreted protein; nutrient; cytokine modulator; proliferation;
XX differentiation; immune system modulator; tissue growth; chemotactic;
XX haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
XX haematopoiesis.
XX
XX OS Homo sapiens.
XX
XX PN WO20011988-A1.
XX
XX 22-MAR-2001.
XX
XX 14-SEP-2000: 2000W0-US25135.
XX
XX 17-SEP-1999: 99US-0394829.

```

(GEM) GENETICS INST INC.

Jacobs K, McCoy JM, LaVallie ER, Collins Racie LA, Evans C;  
 Merberg D, Treacy M, Bowman MR, Spaulding V, Acostino ML;  
 WPI: 2001-244801/25.  
 N-PSDB: AAF98469.

Isolated nucleic acids encoding polypeptides, useful for modulating  
 e.g. cytokine and cell proliferation/differentiation activity, the  
 immune system and hematopoiesis regulating activity -

Disclosure: Page 487-490; 557pp; English.

Human cDNA clones represented in AAF98374 - AAF98489 encode secreted  
 proteins AAB90667 - AAB90750. The cDNA clones are isolated from various  
 tissue types, and may be used in the prevention, treatment and diagnosis  
 of diseases associated with inappropriate protein expression. The  
 polypeptides and nucleic acids may be used as nutrients or to modulate  
 cytokine and cell proliferation/differentiation activity and may also be  
 involved in modulation of the immune system. The cDNA sequences,  
 proteins, their agonists and/or antagonists exhibit hematopoiesis,  
 regulating activity, tissue growth activity, activin/inhibin activity,  
 chemoattractant/chemokinetic activity, haemostatic and thrombolytic  
 activity, receptor/ligand activity, anti-inflammatory activity,  
 haematopoiesis activity, cadherin/tumour suppressor activity, and/or  
 tumour inhibition activity. Included in the invention are probes  
 represented in AAF98490 - AAF98572 which are specific for the cDNA clones  
 encoding the secreted proteins.

XX Sequence 975 AA;  
 SQ

Query Match 100.0%; Score 376; DB 21; Length 975;  
 Best Local Similarity 100.0%; Pred. No. 1,96-34;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLRSGLKRTPSLKPVPPKPSFAP 60  
 DB 904 PPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLRSGLKRTPSLKPVPPKPSFAP 963  
 QY 61 LSTSMKPNDACT 72  
 DB 954 LSTSMKPNDACT 975

RESULT 10  
 ID AAY71460 standard; Protein: 1030 AA.  
 AC AAY71460;  
 DT 04-OCT-2000 (first entry)  
 XX Human semaphorin 6A-1.  
 DE Human semaphorin 6A-1; (HSA)SEMA6A-1; neuronal development;  
 KW neuronal regeneration; Eno/VASP protein family; immunomodulatory;  
 KW gene therapy; diagnostic agent; therapeutic agent; differentiation;  
 KW cytoskeletal stabilisation; plasticity.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 FH Binding-site 959..1030 Location/Qualifiers  
 FT /note= "Zyxin-like domain that selectively binds to  
 members of Eno/VASP protein family, especially Evi"  
 FT Binding-site 957..961  
 FT /note= "Specific binding motif for members of  
 Eno/VASP protein family, especially Evi"  
 FT Binding-site 1009..1014  
 FT /note= "Specific binding motif for members of  
 Eno/VASP protein family, especially Evi"

PN W0200031252-A1.  
 XX  
 PD 02-JUN-2000.  
 XX  
 FF 26-NOV-1999; 99WD-EP0215.  
 XX  
 PR 26-NOV-1998; 98EP-0122441.  
 XX  
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 XX  
 PI Beitz C, Kiostermann A;  
 XX  
 PS WPI: 2000-400065/34.  
 DR N-PSDB: AAD01233.  
 XX  
 PT Nucleic acid coding for human semaphorin 6A-1 used as diagnostic agent,  
 PT therapeutic agent, for modulating immune system, in gene therapy or for  
 PT effecting differentiation, cytoskeletal stabilization and/or plasticity  
 XX  
 XX Example 1; Page 29-33; 5app; English.  
 PS The present sequence is a transmembranous human semaphorin  
 CC 6A-1 (HSA)SEMA6A-1; which is involved in neuronal development and  
 CC regeneration mechanisms during apoptosis. Semaphorin is a family of  
 CC proteins displaying secreted or transmembrane-based repulsive guidance  
 CC cues critically involved in neuronal development. The present sequence  
 CC was isolated from human 1-2AP Express cDNA library which was screened  
 CC using a PCR fragment amplified from human neuroblastoma cell line  
 CC SK-N-MC cDNA. The (HSA)SEMA6A-1 protein contains a zyxin-like domain  
 CC that selectively binds to members of Eno/VASP protein family especially  
 CC Evi. Expression of (HSA)SEMA6A-1 is highest in embryonic brain and  
 CC kidney and moderate in lung. The present sequence is useful as diagnostic  
 CC and therapeutic agents, for modulating the immune system, in gene  
 CC therapy, for effecting differentiation, cytoskeletal stabilisation  
 CC and plasticity.  
 XX  
 SQ Sequence 1030 AA;  
 Query Match 100.0%; Score 376; DB 21; Length 1030;  
 Best Local Similarity 100.0%; Pred. No. 2e-33;  
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLRSGLKRTPSLKPVPPKPSFAP 60  
 DB 959 PPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLRSGLKRTPSLKPVPPKPSFAP 1016  
 QY 61 LSTSMKPNDACT 72  
 DB 1019 LSTSMKPNDACT 1030

RESULT 11  
 ID AAB94239 standard; Protein: 451 AA.  
 AC AAB94239;  
 XX  
 DT 26-JUN-2001 (first entry)  
 DE Human: protein sequence SEQ ID NO:14623.  
 XX  
 KW Human: primer; detection: diagnosis; antisense therapy; gene therapy.  
 XX Homo sapiens.  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-011626.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.

```

PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-018776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Suiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318743/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs.
XX
XX Claim 8: SEQ ID 14623; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 451 AA:
XX
XX Query Match 43.5%; Score 163.5; DB 22; Length 451;
XX Best Local Similarity 50.7%; Pred. No. 5.9e-10;
XX Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;
XX
XX 1 PPAPQRVDSIQVHSSQPSQAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDVPP 54
XX 380 PTPGAKVDYIQ-----GTPSVVHLQPSLSRQSSYTSNGTLPRTGLKRTPSLKPDVPP 432
XX
XX 55 KPSEFAPLSTSMKP 67
XX 433 KPSEFVQTPSVRP 445
XX
XX
XX RESULT 12
XX ID AAH94296
XX AAH94296 standard; Protein: 464 AA.
XX
XX AA594296:
XX
XX 26-JUN-2001 (first entry)
XX
XX Human protein sequence SEQ ID NO:14749.
XX
XX Human: primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX

```

```

1D 07-FEB-2001.
XX
XX 28-JUL-2000; 2000JP-0116126.
XX
XX 29-JUL-1999; 99JP-0248035.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-018776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Suiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318743/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs.
XX
XX Claim 6: SEQ ID 14749; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 464 AA:
XX
XX Query Match 43.5%; Score 163.5; DB 22; Length 464;
XX Best Local Similarity 50.7%; Pred. No. 6.1e-10;
XX Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;
XX
XX 1 PPAPQRVDSIQVHSSQPSQAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDVPP 54
XX 393 PTPGAKVDYIQ-----GTPSVVHLQPSLSRQSSYTSNGTLPRTGLKRTPSLKPDVPP 445
XX
XX 55 KPSEFAPLSTSMKP 67
XX 446 KPSEFVQTPSVRP 458
XX
XX
XX RESULT 13
XX ID ABU11724
XX ABU11724 standard; Protein: 474 AA.
XX
XX ABU11724:
XX
XX 13-FEB-2003 (first entry)
XX
XX Human MCDT polypeptide SEQ ID 671.
XX
XX MCDT; human; disease detection and treatment molecule polypeptide;
XX

```

KW	anti-inflammatory; immunosuppressive; osteopathic; cystostatic; anti-HAV;
KW	haemostatic; neptrotropic; antitumoric; antiparasitic; hepatotropic;
KW	gene therapy; protein replacement therapy; cell proliferative disorder;
KW	cancer; aconocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma;
KW	anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
KW	Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
KW	psoriasis; hepatitis.
XX	
OS	Homo sapiens.
XX	
PN	WC020279449-A2.
PD	10-OCT-2002.
XX	
PF	27-MAR-2002; 2002WC-US09944.
XX	
PR	28-MAR-2001; 2001US-279619P.
PR	29-MAR-2001; 2001US-280067P.
PR	29-MAR-2001; 2001US-280068P.
PR	16-MAY-2001; 2001US-291280P.
PR	17-MAY-2001; 2001US-291829P.
PR	17-MAY-2001; 2001US-291849P.
PR	19-JUN-2001; 2001US-293428P.
PR	20-JUN-2001; 2001US-293776P.
PR	20-JUN-2001; 2001US-30000-P.
XX	
PA	{INCY-} INCYTE GENOMICS INC.
XX	
PI	Daffo A, Jones AL, Tran AE, Dahi CR, Gietzen D, Chien J;
PI	Duford GE, Hillman JL, Yu JY, Tusson O, Yap PE, Austey SK;
PI	Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gestin EE;
PI	Peralta CH, David MB, Lewis SA, Chen AC, Packer SP, Harris B;
PI	Flores V, Marwaha R, Lo A, Ian RV, Uraskin ME;
DR	WPI: 2003-058431/05.
DR	N-PSDB: ABX34714.
XX	
PT	New purified disease detection and treatment molecule proteins and
PT	polynucleotides, useful for diagnosing, treating or preventing cancers
PT	(e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis
PI	or hepatitis .
PI	
PS	Claim 27: SEQ ID NO 671: 33pp + Sequence listing: English.
XX	
CC	This invention describes a novel disease detection and treatment molecule
CC	polypeptide (MDD) which has anti-inflammatory, immunosuppressive,
CC	osteopathic, cystostatic, anti-HIV, haemostatic, hepatotropic,
CC	antitumoric, antiparasitic and hepatotropic activity. The polynucleotides
CC	and the polypeptides of the invention can be used for gene therapy of
CC	protein replacement therapy and are useful for treating a variety of
CC	diseases or conditions. These polypeptides or polynucleotides are
CC	particularly useful for diagnosing, treating or preventing cell
CC	proliferative disorders (e.g. cancers including adenocarcinoma,
CC	leukemia, lymphoma, melanoma, myeloma or sarcoma), anemia, Crohn's
CC	disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's
CC	syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or
CC	hepatitis. AB011450-AB011845 represent the MDT polynucleotides encoded
CC	by AB011450-AB011845, described in the disclosure of the invention.
CC	NOTE: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format from Wipo at
XX	ftp.wipo.int/pub/published_pct_sequences.
XX	
SO	Sequence 474 AA:
	Query Match 43.5%; Score 163.5; DB 24: Length 474;
	Best Local Similarity 50.7%; Pred. No. 6,3e-10;
	Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2.
OY	1 PPPAPCRVDSIOVHSSOPS30AAVVSROP:SLINAYNSLIT-----RSLKRTSLKPDVPP 54
Db	403 PPTTCAKVDYTO-----GTPSVHLQPSLSROSSYTSNGLPLRGLKRPISLKPDPVP 455
OY	55 KPSFAPLSTSKKP 67







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2003, 16:22:28 ; Search time 44 seconds  
(without alignments)  
30,347 Million cells / pairs of seq

Title: US-09-856-681-4

Perfect score: 376

Sequence: 1 PPAPQVDSIQVHSSQPSQ.....PDKSAPLSTSMKPLAEE 12

Scoring table: BLOSUM62

Gapop 10.0 / Gapext 0.5

Searched: 283308 seqs, 9616882 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

PTR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	87	23.1	961	2 A55360	faciogenital dyspl
2	80.5	21.4	1322	2 A50288	myosin heavy chain
3	77.5	20.6	1111	2 I05646	hypothetical prote
4	77	20.5	175	2 I47453	serine/proline-ric
5	76.5	20.3	744	2 E86255	hypothetical prote
6	75.5	20.1	393	2 I33103	fib-1 protein - Ca
7	75.5	20.1	494	2 A42170	zinc finger protei
8	75.5	20.1	457	2 J05076	myo-associated zin
9	74	19.7	452	2 S22159	imidazoleglycerol
10	73.5	19.5	282	2 I42717	CNA-binding protei
11	73	19.4	867	2 I41309	hypothetical zinc-
12	72.5	19.3	628	2 S01955	hypothetical prote
13	72.5	19.3	457	2 S84869	probable SRF6 prot
14	72.5	19.3	4357	2 T03455	ALK protein - huma
15	72.5	19.3	5262	2 T03454	ALK protein - huma
16	72	19.1	459	2 A41277	retinoid acid rece
17	71.5	19.0	446	2 A42229	transcription fact
18	71.5	19.0	1522	2 I35471	transcription fact
19	71.5	19.0	2957	2 I35152	hypothetical prote
20	71	18.9	621	2 J07278	adaptor protein co
21	71	18.9	1150	2 S56775	myel protein - stru
22	70.5	18.8	468	2 I46815	hypothetical prote
23	70.5	18.8	1420	2 I35781	probable cytoskele
24	70	18.6	719	2 S62466	ATP-dependent RNA
25	70	18.6	747	2 S35546	hypothetical prote
26	70	18.6	792	2 T26050	hypothetical prote
27	70	18.6	1012	2 I53172	PAE-28 - mouse
28	70	18.6	1201	2 G85441	unknown protein li
29	69.5	18.5	331	2 B47236	zinc-finger protei

30 59.5 18.5 477 2 A47236 zinc-finger protei  
31 59.5 18.5 125 2 B41206 microtubule-associ  
32 69 18.4 3942 2 T42730 Bassoon protein -  
33 68.5 18.2 623 2 S48941 regulatory protein  
34 68.5 18.2 1106 2 T31742 hypothetical prote  
35 68.5 18.2 1189 2 S49915 extensin-like prot  
36 68 18.1 428 2 TV008K transforming prote  
37 68 18.1 533 2 A45690 transactivator EBV  
38 68 18.1 613 2 T47975 auxin response fac  
39 68 18.1 736 2 T25447 hypothetical prote  
40 68 18.1 983 2 T40873 probable transcrip  
41 68 18.1 1172 2 T00065 hypothetical prote  
42 68 18.1 1215 2 I61713 co-repressor prote  
43 68 18.1 1226 2 A56068 co-repressor prote  
44 68 18.1 1238 2 J05765 histonol polyphosp  
45 67.5 18.0 429 2 J04965 euk1 protein - mou

#### ALIGNMENTS

##### RESULT 1

A55360

faciogenital dysplasia-associated protein FGS1 - human

C:Species: Homo sapiens (man)

C:Date: 10-Feb-1995 #sequence\_revision 10-Feb-1995 #text\_change 17-Mar-1999

C:Accession: A55360

K:Posteris: N.G.; Cadle, A.; Lajler, L.J.; Porteous, M.E.M.; Schwartz, C.E.; Stevens  
Cell 79, 669-676, 1994

A:Title: Isolation and characterization of the faciogenital dysplasia (Aarskog-Scott

A:Reference number: A55360; MUID:95042764; PMID:7954831

A:Accession: A55360

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-961 <PAS>

A:Cross-references: GB:011690; NID:G595424; PID:G595425

C:Superfamily: CDC24 homology; p10ckstrin repeat homology

P:573-561/Domain: CDC24 homology <CD24>

Query Match 23.1% Score 87; DB 2; Length 961;

Best Local Similarity 34.8%; Pred. No. 1.2;

Matches 23; Conservative 5; Mismatches 22; Indels 16; Gaps 2;

QY 2 PPAPQVDSIQVHSSQPSQAVTVSRQPSLNAYNSLTRSGIKRTPSLKQVPPKPSAPL 61

DB 127 PEGFORL-----KSDGQFPETPSQRP-----SPLKRAPGKQVPPKPSYLM 170

QY 52 STSMKP 67

DB 171 PRMPPE 176

##### RESULT 2

A59288

myosin heavy chain Myr 8 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 08-Sep-2000

C:Accession: A59288

R:Patel, K.G.; Liu, C.; Cameron, P.L.; Cameron, R.S.

submitted to GenBank, November 1999

A:Description: Identification of a Novel Mammalian Myosin Class, XVI, in Developing

A:Reference number: A59288

A:Accession: A59288

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1322 <PAI>

A:Cross-references: GB:A59288114; PID:A59288114

A:Experimental source: strain Sprague-Dawley; clone KP4; cell type type 1 astrocyte

C:Superfamily: myosin motor domain homology

F:404-1132/Domain: myosin motor domain homology <MCO>

Query Match 21.4%; Score 80.5; DB 2; Length 1322;

Best Local Similarity 38.4%; Pred. No. 7.4;







## RESULT 15

T03454

ALR protein - human

C:Species: Homo sapiens (man)

C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_exchange 21-Jul-2003

C:Accession: T03454

R:Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, P.; Bruck, T.; Ballupally, R.; Yanov,

Oncogene 15, 549-560, 1997.

A:Title: Structure and expression pattern of human ALR, a novel gene with strong homology

A:Reference number: Z14954; E01357386474; E0101421308

A:Accession: T03454

A:Status: preliminary; translated from G5/EMBL/D047

A:Molecule type: mRNA

A:Residues: 1-5262 &lt;PRA&gt;

A:Cross-references: EMBL:AF010403; NID:q2356234; PIDD:AAC51734.1; FID:q2356285

C:Genetics:

A:Gene: ALR

A:Map position: 12

C:Superfamily: human ALR protein

C:Keywords: alternative splicing

## Query Match

Db 12 QVHSSQPSGAVTSVPSNVAVNSITRSGLKX-----LSLRDYP---PKP --- 56

Best local similarity 34.7% Prod. No. 201023

Matches 25 Conservative 6; Mismatches 24; Indels 17; Gaps 4;

QY 12 QVHSSQPSGAVTSVPSNVAVNSITRSGLKX-----LSLRDYP---PKP --- 56

Db 2230 ELKAKVPSGPPNPNVRSCTGTGAPVG-TTSPMRFTFOAVGFPSLKPVPQPSLPDPPHCIN LDR

QY 57 -SPAP-STSMPK 67

Db 2289 SHFPGFTLGKP 2300

QY 57 -SPAP-STSMPK 67

Db 2289 SHFPGFTLGKP 2300

QY 57 -SPAP-STSMPK 67

Db 2289 SHFPGFTLGKP 2300

QY 57 -SPAP-STSMPK 67

Db 2289 SHFPGFTLGKP 2300

Search completed: September 30, 2003, 16:35:30

Job time : 55 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw mode:

Run on: September 30, 2003, 15:15:02 : Search time 38 Seconds  
(without alignment) 34 bytes/sec  
69,103 Million cell updates/sec

Title: US-09-856-681-4

Perfect score: 376

Sequence: 1 PPAFQVDSIQVHSQSG.....FKKSEAPLSTSMKNDACT 72

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127653 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127653

Minimum DB seq length: 0

Maximum DB seq length: 2600000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 4% summaries

Database : SwissProt\_41\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	376	100.0	1030	1	SMGA_HUMAN
2	87	23.1	961	1	FGD1_HUMAN
3	86.5	23.0	564	1	2YX_MOUSE
4	86	22.9	862	1	MAK3_RAT
5	77	20.5	960	1	FGD1_MOUSE
6	75.5	20.1	397	1	GAT5_HUMAN
7	75.5	20.1	477	1	MAZ_HUMAN
8	75.5	20.1	5147	1	PCO2_HUMAN
9	74	19.7	452	1	HIS7_HYPR
10	74	19.7	894	1	MAK3_HUMAN
11	73.5	19.5	5085	1	PCO2_RAT
12	72.5	19.3	628	1	V7OK_TYRV
13	71.5	19.0	446	1	TEB3_MOUSE
14	71.5	19.0	669	1	SEEL_HUMAN
15	71.5	19.0	1004	1	PHC1_HUMAN
16	71.5	19.0	1522	1	PS1L_SCHRO
17	71	18.9	428	1	ELK1_HUMAN
18	70.5	18.8	344	1	Z1PA_SHEON
19	70	18.6	719	1	PRK1_SCHRO
20	70	18.6	1012	1	PHC1_MOUSE
21	69.5	18.5	331	1	MAZ_MOUSE
22	69.5	18.5	477	1	MAZ_MOUSE
23	69.5	18.5	1125	1	MAP4_MOUSE
24	58.5	18.2	625	1	R1Q1_YEAS
25	58.5	18.2	743	1	TEB3_HUMAN
26	68	18.1	737	1	SKN1_CANAL
27	68	18.1	812	1	NAH2_HUMAN
28	68	18.1	1259	1	AU12_HUMAN
29	67.5	18.0	429	1	ELK1_MOUSE
30	67.5	18.0	525	1	CO2A_HUMAN
31	67.5	18.0	629	1	V7OK_TYMYA
32	67.5	18.0	629	1	V7OK_TYMYC
33	67.5	18.0	5179	1	MUC21_HUMAN

#### ALIGNMENTS

##### RESULT 1

ID	SMGA_HUMAN	STANDARD	PRT: 1030 AA
AC	Q9H2E6: Q9H2E6		
BI	15-SEP-2003 (rel. 42, Created)		
DI	15-SEP-2003 (rel. 42, Last sequence update)		
FI	15-SEP-2003 (rel. 42, Last annotation update)		
IS	Semaphorin 6A precursor (Semaphorin VIA) (Sema VIA) (Semaphorin 6A-1)		
LS	(SPM6A-1)		
OS	SEMATA OR KIAA1143		
CS	Homo sapiens (Human)		
CC	Eukaryota; Metazoa; Chordata; Vertebrata; Puteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID:9606		
RN	[1]		
RP	SEQUENCE FROM N.A., AND INTERACTION WITH EVL.		
RA	MEDLINE:20564339; PubMed 10993894;		
EX	Klostermann A., Lutz H., Sertler F., Beil G.: "The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1/Sema6A-1) bind to the enabled/vasodilator-stimulated phosphoprotein-like protein (EVL) via a novel carboxyl-terminal zyxin-like domain."		
RT	J. Biol. Chem. 275:35647-35653(2000).		
RL	[2]		
RN	SEQUENCE FROM N.A.		
RP	TISSUE:Brain;		
RC	MEDLINE:20181126; PubMed-10718198;		
RA	Nakase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.: "Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro."		
RT	DNA Res. 7:65-73(2000).		
RL	[1]		
CC	FUNCTION: Can act as repulsive axon guidance cues. May play a role in channeling sympathetic axons into the sympathetic chains and controlling the temporal sequence of sympathetic target innervation (By similarity).		
CC	SUBUNIT: Active as a homodimer or oligomer. Interacts with EVL.		
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.		
CC	-1- ALTERNATIVE PRODUCTS:		
CC	Event=Alternative splicing; Named isoforms=2;		
CC	Name=1;		
CC	Isoid-Q9H2E6-1; Sequence=Displayed;		
CC	Name=2;		
CC	Isoid-Q9H2E6-2; Sequence=VSP_007113;		
CC	Note=NO EXPERIMENTAL CONFIRMATION AVAILABLE;		
CC	-1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.		
CC	-1- SIMILARITY: Contains 1 Sema domain.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	-----		

34 67 17.8 315 1 YK04\_CAFEL  
35 67 17.8 529 1 DNB2\_ADE05  
36 67 17.8 538 1 KNC0\_YEAST  
37 67 17.8 779 1 SRP\_PROMO  
38 67 17.8 813 1 NAH2\_RAT  
39 67 17.8 1078 1 S24A\_HUMAN  
40 67 17.8 1152 1 NSP4\_HUMAN  
41 67 17.8 2035 1 HEC1\_HUMAN  
42 66.5 17.7 563 1 WAP\_HUMAN  
43 66.5 17.7 837 1 ROL1\_YEAST  
44 66 17.6 295 1 PF1B\_MYCLE  
45 66 17.6 433 1 HXB3\_MOUSE







between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: AF12224; AAK532.4.1; -  
 HSSP: P24941; 1WHH  
 InterPro: IPR000180; GifPro.  
 InterPro: IPR000719; Pro\_Kinase.  
 InterPro: IPR002290; Ser\_thr\_kinase.  
 Pfam: PF00780; CNH; 1.  
 Pfam: PF00069; pkinase; 1.  
 ProDom: PD003001; Prot\_kinase; 1.  
 SMART: SM00036; CNH; 1.  
 SMART: SM00320; S\_TKc; 1.  
 PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 PROSITE: PS00108; PROTEIN\_KINASE\_ST; FALSE\_NEG.  
 ATP-binding: Transferrase; Serine/threonine-protein kinase.  
 NON\_TER 1  
 DOMAIN 5 262 PROTEIN KINASE.  
 FT DOMAIN 530 842 CNH.  
 FT NP\_BIND 11 13 ATP (BY SIMILARITY).  
 FT BINDING 34 34 ATP (BY SIMILARITY).  
 FT ACT\_SITE 125 125 BY SIMILARITY.  
 FT ACT\_SITE 125 125 BY SIMILARITY.  
 SQ SEQUENCE 862 AA: 97390 MW: 58013AC3B0A3287F CRC64;  
 Query Match 22.9% Score 86; DB 1; Length 862;  
 Best Local Similarity 34.7%; Pred. No. 0.72;  
 Matches 25; Conservative 5; Mismatches 27; Indels 14; Gaps 3;  
 QY : PPPAPQVDS:TCV-----HSSQFSGQAVTSKQSLNAYNSLTSGLKRTFSIKHWPEK 56  
 DB 400 PPLPAPKSTSIPODTHSHSDSNQSG:TKKCS-----SSPAKSHVPPPPPP 449  
 QY 57 STAPLSTSMKPN 58  
 DB 450 RLPPQKPAVIGN 46;  
 RESULT 5  
 FGD1\_MOUSE  
 ID FGD1\_MOUSE STANDARD: P87 960 AA.  
 AC P52734;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 26-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative rho/rac guanine nucleotide exchange factor (Rho/Rac GEF)  
 DE (Pleioquinatal dysplasia protein homolog).  
 GN FGD1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=96081343; PubMed=8538075;  
 RX Pasteris N.G., de Gouyon B., Cadie A.R., Campbell K., Horvath G.E.,  
 RA Gorski J.;  
 RT Cloning and regional localization of the mouse pleioquinatal  
 RT dysplasia (Pqdl) gene.  
 RL Mamm. Genome 6:558-561(1995).  
 CC -!- FUNCTION: ACTIVATES THE RAS-LIKE FAMILY OF RHO AND RAC PROTEINS  
 CC BY EXCHANGING BOUND GDP FOR FREE GTP.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Contains 1 Dbl-homology (DB) domain.  
 CC -!- SIMILARITY: Contains 2 FVE-Grecks.  
 CC -!- SIMILARITY: Contains 1 FVE-type zinc finger.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: U22325; AAK56001.1; -  
 HSSP: Q07889; 1AWH  
 MCD: MCD104566; Pqd1.  
 InterPro: IPR001331; GDS\_GDC24.  
 InterPro: IPR001849; PH.  
 InterPro: IPR000219; RhoGEF.  
 InterPro: IPR000306; Znf\_FYVE.  
 Pfam: PF01363; FYVE; 1.  
 Pfam: PF00169; PH; 2.  
 Pfam: PF00621; RhoGEF; 1.  
 SMART: SM00064; FYVE; 1.  
 SMART: SM00233; PH; 2.  
 SMART: SM00325; RhoGEF; 1.  
 PROSITE: PS00741; DH; 1; FALSE\_NEG.  
 PROSITE: PS00010; DH; 2; 1.  
 PROSITE: PS00063; PH\_DOMAIN; 2.  
 PROSITE: PS0178; ZF\_FYVE; 1.  
 KW Guanine-nucleotide releasing factor; Zinc-finger; Repeat.  
 FT DOMAIN 372 560 PH.  
 FT DOMAIN 7 330 PRO-RICH.  
 FT SLIE 171 179 SH3-BINDING (POTENTIAL).  
 FT SLIE 179 187 SH3-BINDING (POTENTIAL).  
 FT DOMAIN 589 688 PH 1.  
 FT ZNF\_FING 729 789 FYVE-TYPE.  
 FT DOMAIN 820 920 PH 2.  
 SQ SEQUENCE 960 AA: 106477 MW: 4C1B84DF490FC51 CRC64;  
 Query Match 20.5% Score 77; DB 1; Length 960;  
 Best Local Similarity 31.8%; Pred. No. 5.9;  
 Matches 21; Conservative 5; Mismatches 24; Indels 16; Gaps 2;  
 QY 2 PPAPQVDS:TCV-----HSSQFSGQAVTSKQSLNAYNSLTSGLKRTFSIKHWPEK 61  
 DB 129 PEGPQF-----KSDPGFPIEIPGRP-----SPLKRAPGPKPQVPPKPSYLM 170  
 QY 62 SLSMKP 67  
 DB 271 PVIIPP 176  
 RESULT 6  
 GAT5\_HUMAN  
 ID GAT5\_HUMAN STANDARD: P87 397 AA.  
 AC QYHXX5;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 26-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Transcription factor GATA-5 (GATA binding factor-5).  
 GN GATA5  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=21638749; PubMed=11780052;  
 RA DeCausas P., Matthews D.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baquley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley C.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Gratham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,

RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Ray M.P., Kimberley A.M., King A., Kligfus A., Laird G.K., Lawler S.,  
RA Leivaestaino M.H., Loversha M.A., Lloyd G., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson J.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.L.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund G., Steward C.A., Sillston J.E.,  
RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Treman A.C., Vaudin R., Vee M., Williams E.M.,  
RA Whitehead S.L., Whitaker P., Willey D., Williams E., Williams S.A.,  
RA Wilming L., Wray P.W., Richard T., Gartin R.M., Bentley D.R., Beck S.,  
RA Rogers J.,  
RT "The DNA sequence and comparative analysis of human chromosomes 20-22  
RL Nature 414:655-671 (2001).  
CC -1- FUNCTION: BINDS TO THE FUNCTIONALLY IMPORTANT DEF-1 NUCLEAR  
CC PROTEIN BINDING SITE IN THE CARDIAC-SPECIFIC SLOW/CARDIAC TROPONIN  
CC C TRANSCRIPTIONAL ENHANCER. MAY PLAY AN IMPORTANT ROLE IN THE  
CC TRANSCRIPTIONAL PROGRAM(S) THAT UNDERLIES SMOOTH MUSCLE CELL  
CC DIVERSITY (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC -1- SIMILARITY: Contains 2 C2H2-type zinc fingers.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: A249627; 1CNF; 1.  
CC HSSP: P17679; 1CNF; 1.  
CC GENE: HGNC:15802; GATA5.  
CC InterPro: IPR006679; Znf\_GATA.  
CC Pfam: PF05349; GATA-N; 1.  
CC Pfam: PF00320; GATA; 2.  
CC PRINTS: PR00619; GATAZNFINGER.  
CC SMART: SM00401; 2.F.GATA; 2.  
CC PROSITE: PS00344; GATA\_ZNFINGER\_1; 2.  
CC PROSITE: PS01114; GATA\_ZNFINGER\_2; 2.  
CC Transcription: regulation; Activator; DNA-binding; Zinc-finger;  
CC Nuclear protein.  
CC FT 2N.FING 189 213 GATA-TYPE.  
CC FT 2N.FING 243 267 GATA-TYPE.  
CC SQ SEQUENCE 397 AA: 41299 MW: 50FBA0208695C57 CRK64:  
Query Match 20.1% Score 75.5; DB 1; Length 397;  
Best Local Similarity 29.1%; Pred. No. 3;  
Matches 22; Conservative 6; Mismatches 36; Indels 5; Gaps 1;  
QY 1 PPAPORVUSIQVHSSQP-----SQAVTVSPQPSLNAYNSLRGLKRRLVSLKPL 51  
DB 277 PRPLAMKKEIQIRKKPKPIAKRGSSGSTRNASASPSAVASTOSAAATSKAKESIASP 124  
QY 52 VPKPSFAPLSTSMK 65  
DB 337 VDPGRSMARQASQCE 351  
RESULT 7  
MAZ\_HUMAN  
ID MAZ\_HUMAN STANDARD; PRT: 477 AA.  
AC P56270; Q15763; Q99443;  
DT 15-JUL-1993 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE MYC-associated zinc finger protein (MAZI) (Purine binding  
DE transcription factor) (Pur-1) (ZF87) (ZNF87).  
GN MAZ.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; CranLata; Vertebrata; Euteleostomi;

CC Mammalia; Butteria; Primates; Catarrhini; Hominoidea; Homo.  
CC NCBI\_TaxID:9606;  
CC -----  
CC SEQUENCE FROM N.A.  
CC MEDLINE:92366475; PubMed:3502157;  
CC Bressane S.A., Asselin C., Patel A.J., Marcu K.B.,  
CC "MAZ, a zinc finger protein, binds to c-MYC and C2 gene sequences  
CC regulating transcriptional initiation and termination.",  
CC Proc. Natl. Acad. Sci. U.S.A. 89:7452-7456(1992).  
CC -----  
CC SEQUENCE FROM N.A.  
CC MEDLINE:92323279; PubMed:1567856;  
CC Pyic G.J., Motera K.H., Hall D.J.,  
CC "Isolation of a novel cDNA encoding a zinc-finger protein that binds  
CC to two sites within the c-myc promoter.",  
CC Biochemistry 31:4102-4110(1992).  
CC -----  
CC SEQUENCE FROM N.A.  
CC TISSUE:Pancreatic islets;  
CC MEDLINE:9642551; PubMed:8431693;  
CC Tsutsui H., Sakatsune O., Itakura K., Yokoyama K.K.,  
CC "Members of the MAZ family: a novel cDNA clone for MAZ from human  
CC pancreatic islet cells.",  
CC Biochem. Biophys. Res. Commun. 226:801-809(1995).  
CC -----  
CC SEQUENCE FROM N.A.  
CC MEDLINE:96224025; PubMed:4626793;  
CC Parks C.L., Shenk T.,  
CC "The serotonin 1a receptor gene contains a TATA-less promoter that  
CC responds to MAZ and Sp1.",  
CC J. Biol. Chem. 271:4417-4430(1996).  
CC -----  
CC SEQUENCE FROM N.A.  
CC TISSUE:Myophoblastoma;  
CC MEDLINE:98352105; PubMed:9685418;  
CC Song J., Kurakami H., Tsutsui H., Iano X., Matsumura M., Itakura K.,  
CC Kanazawa I., Sun K., Yokoyama K.K.,  
CC "Genomic organization and expression of a human gene for Myc-  
CC associated zinc finger protein (MAZ).",  
CC J. Biol. Chem. 273:20603-20614(1998).  
CC -----  
CC FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR WITH DUAL ROLES  
CC IN TRANSCRIPTION INITIATION AND TERMINATION. BINDS TO TWO SITES,  
CC MET1 AND MET2, WITHIN THE C-MYC PROMOTER HAVING GREATER  
CC AFFINITY FOR THE FORMER. ALSO BINDS TO MULTIPLE G/C-RICH SITES  
CC WITHIN THE PROMOTER OF THE SPI FAMILY OF TRANSCRIPTION FACTORS.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- TISSUE SPECIFICITY: HEART, BRAIN, PLACENTA, LUNG, LIVER, SKELETAL  
CC MUSCLE, AND PANCREAS. SEEMS NOT TO BE EXPRESSED IN KIDNEY.  
CC -1- SIMILARITY: Contains 6 C2H2-type zinc fingers.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: M94046; 1; NOT\_ANNOTATED\_CDS.  
CC EMBL: M93339; 1; NOT\_ANNOTATED\_CDS.  
CC EMBL: D85131; BAA12728.1; ALT\_INIT.  
CC EMBL: U33619; AAB04121.1; ALT\_INIT.  
CC EMBL: AB017335; BAA33064.1; 1.  
CC PIR: A42170; A42170.  
CC TRANSFAC: T00490; 1.  
CC TRANSFAC: T02105; 1.  
CC GENE: HGNC:6914; MAZ.  
CC MIM: 600999; 1.  
CC GO: 0005367; P:transcription initiation from Pol II promoter; TAS.  
CC GO: 0006369; P:transcription termination from Pol II promoter; TAS.  
CC InterPro: IPR007087; Znf\_C2H2.  
CC Pfam: PF00096; Zf-C2H2; 5.



```

DR InterPro: IPR000008; C2.
DR InterPro: IPR001565; Synaptotagmin.
DR PRINTS: PR03369; C2DOMAIN.
DR PRINTS: PR00399; SYNAPTOTAGMIN.
DR SMART: SMO0239; C2; 2.
DR PROSITE: PS00499; C2DOMAIN.1; 1.
DR PROSITE: PS00004; C2DOMAIN.2; 2.
KW Calcium/phospholipid-binding; Zinc; Metal-binding; Zinc-finger;
KW Repeat; Alternative splicing.
FT NON_TER 1
FT DOMAIN 400 465
FT 2N_FING 499 523
FT 2N_FING 969 992
FT NON_CONS 1010 1011
FT DOMAIN 2300 2325
FT DOMAIN 4391 4442
FT DOMAIN 4544 4633
FT DOMAIN 5031 5121
FT VARSPLIC 4404 4404
FT 10 X 10 AA TANDEM APPROXIMATE REPEATS OF
P-A-K-P-Q-P-Q-P-X.
C4-TYPE (POTENTIAL).
C4-TYPE (POTENTIAL).
POLY-PRO.
C2 DOMAIN 1.
C2 DOMAIN 2.
S -> SCNGLGRIVGKELRHSQELDAYIAKIIKGSAAE
QICKMES (in isoform 2).
/FTID-VSP_003924.
X -> KPTDGRKVSHTFTRD (in isoform 2).
/FTID-VSP_003924.
G -> GQVNVONAS (in isoform 2).
/FTID-VSP_003924.
TAHKS -> SKRRK (in isoform 2).
/FTID-VSP_003924.
MISSING (in isoform 2).
/FTID-VSP_003924.
FT VARSPLIC 4757 4761
FT VARSPLIC 4762 5147
FT SEQUENCE 5147 AA; 564537 MW; 6D5D8499C498C3E0 CRC64;
Query Match 20.1%; Score 25.5; DR 1; Length 5147;
Best Local Similarity 31.4%; Pred. No. 53;
Matches 22; Conservative 5; Mismatches 18; Indels 25; Gaps 4;
QY 1 PPAPFQVDSIOVHSSQKAGQAVTVSRPSINAYNSLTRSLKRTPS:KPVFEPKPSFAP 60
DQ 2378 PPVPPKPSI-----PGLVFTIETPERS-----KPTAPKPVTP 2413
QY 61 J-STSMKPNK 69
DQ 2414 LPTTQKPTD 2423
RESULT 9
HIS7_PHYR
ID HIS7_PHYR STANDARD: PRT: 452 AA.
AC P28624;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19) (IGPD).
GN HIS3.
OS Phytophthora parasitica (Potato buckeye rot agent).
OC Eukaryota; Stramenopiles; Oomycetes; Pythiales; Pythiaceae;
OC Phytophthora.
OX NCBI_TaxID:4792;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN:DSM 1829;
RA Baitrusch-Weiler M., Karlovsky P., Pirel R.H.;
RL Submitted (JAN-1992) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: D-erythro-1-(imidazo[4-y]glycerol 3
CC phosphate = 3-(imidazo[4-y] 2-oxopropyl phosphate - H(2)O.
CC -!- PATHWAY: Histidine biosynthesis; sixth step.
CC -!- SIMILARITY: Belongs to the imidazoleglycerol-phosphate dehydratase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL/GenBank/DBJ
CC European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z11591; CAA77675.1; -.
DR InterPro: IPR006438; HAD-SF-1A-hypl.
DR InterPro: IPR005843; Histidinol-phos.
DR InterPro: IPR005834; Hydrolase.
DR InterPro: IPR000807; IGPD.
DR Pfam: PF00702; Hydrolase; 1.
DR Pfam: PF00475; IGPD; 1.
DR ProDom: PD02282; IGPD; 1.
DR TIGRfam: TIGR01548; HAD-SF-1A-hypl; 1.
DR TIGRfam: TIGR01656; Histidinol-ppas; 1.
DR PROSITE: PS00954; GP-DEHYDRATASE_1; 1.
DR PROSITE: PS00955; GP-DEHYDRATASE_2; 1.
KW Histidine biosynthesis; Lyase; Multi-functional enzyme.
FT DOMAIN 1 233 UNKNOWN ACTIVITY.
FT DOMAIN 234 452 IMIDAZOLEGLYCEROL-PHOSPHATE DEHYDRATASE.
FT SEQUENCE 452 AA; 47661 MW; CAE65BE32A9E53A1 CRC64;
Query Match 19.7%; Score 74; DR 1; Length 452;
Best Local Similarity 34.4%; Pred. No. 49;
Matches 21; Conservative 8; Mismatches 22; Indels 10; Gaps 3;
QY 12 QVHSSQKAGQAVTVSRPSINAYNSLTRSLKRTPS:KPVFEPKPSFAPLSTSM 65
DQ 11 EHRQPKGKXAVTVSR-PRKDKAKPLTHGIE---DLFPVQIWLKDCPKPSPEPILLAL 166
QY 66 K 66
DQ 167 K 167
RESULT 10
M4K3_HUMAN
ID M4K3_HUMAN STANDARD: PRT: 894 AA.
AC O81VH6; O81VH7; O97UM5; Q9Y555;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 3 (EC 2.7.1.37)
DE (MAKK/ERK kinase kinase 3) (MEK kinase kinase 3) (MEKKK 3)
DE (Germinal center kinase related protein kinase) (GLK).
GX M4K3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2), FUNCTION, TISSUE SPECIFICITY, AND
RP MUTAGENESIS OF LYS-48.
RC ISSUE-Macrophage, and Skeletal muscle;
RX MEDLINE-97420743; PubMed-9275185;
RA Diener K., Wang X.S., Chen C., Meyer C.F., Keesler G., Zukowski M.,
RA Tan T.-H., Yao Z.;
RT "Activation of the c-Jun N-terminal kinase pathway by a novel protein
RT kinase related to human germinal center kinase.";
RC Proc. Natl. Acad. Sci. U.S.A. 94:9687-9692(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
RA Gerry M.C., Zhang Y., Marks J.J., Suppe B., Hart S., Cortelli J.,
RA Pallos D., Hart T.C.;
RT "Physical/genetic map of the 2p22-2p21 region on chromosome 2.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-712 FROM N.A. (ISOFORM 1).
RA Edwards J., Wohldmann P., Hawkins M., Harkins R.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May play a role in the response to environmental stress.
CC Appears to act upstream of the c-Jun N-terminal
CC pathway.
CC -!- CATALYTIC ACTIVITY: ATP -> a protein + ADP + a

```

phosphoprotein.  
 -1- CoPACTOR: Mannesim.  
 -1- SUBUNIT: Interacts with SHG2. Interaction appears to require MAPK3-mediated JNK activation (By similarity).  
 -1- ALTERNATIVE PROUCTS:  
 Event-Alternative splicing: Named Isoforms:  
 Name=1:  
 IsoId=Q8IVH8-1; Sequence-Displayed;  
 Note-No experimental confirmation available;  
 Name=2;  
 IsoId=Q8IVH8-2; Sequence-VSP\_007052;  
 Name=3;  
 IsoId=Q8IVH8-3; Sequence-VSP\_007053;  
 Note-No experimental confirmation available;  
 -1- TISSUE SPECIFICITY: Ubiquitously expressed in all tissue examined, with high levels in heart, brain, placenta, skeletal muscle, kidney and pancreas and lower levels in lung and liver.  
 -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. STR20 SUBFAMILY.  
 -1- SIMILARITY: Contains 1 CNH domain.  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AF445400; AAN75850.1; JOINED.  
 DR EMBL: AF445401; AAN75850.1; JOINED.  
 DR EMBL: AF445402; AAN75850.1; JOINED.  
 DR EMBL: AF445403; AAN75850.1; JOINED.  
 DR EMBL: AF445404; AAN75850.1; JOINED.  
 DR EMBL: AF445405; AAN75850.1; JOINED.  
 DR EMBL: AF445406; AAN75850.1; JOINED.  
 DR EMBL: AF445407; AAN75850.1; JOINED.  
 DR EMBL: AF445408; AAN75850.1; JOINED.  
 DR EMBL: AF445409; AAN75850.1; JOINED.  
 DR EMBL: AF445410; AAN75850.1; JOINED.  
 DR EMBL: AF445411; AAN75850.1; JOINED.  
 DR EMBL: AF445412; AAN75850.1; JOINED.  
 DR EMBL: AC007684; AAF19249.1;  
 HSSP: P24941; 1B36.  
 Genew: HGNC:6955; MAP4K3.  
 MIM: 604921;  
 DR InterPro: IPR001180; Citron.  
 DR InterPro: IPR000719; Prot\_Kinase.  
 DR InterPro: IPR002230; Ser\_thr\_kinase.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR Pfam: PF00780; CNH; 1.  
 DR Pfam: PF00669; pkinase; 1.  
 DR ProDom: PD000001; Prot\_Kinase; 1.  
 DR SMART: SM00036; CNH; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_SI; FALSE\_NEG.  
 KW ATP-binding; transferase; Serine/threonine-protein kinase;  
 KW Alternative splicing.  
 FT DOMAIN 16 273 PROTEIN KINASE.  
 FT DOMAIN 562 874 CNH.  
 FT NP\_S'ND 22 30 ATP (BY SIMILARITY).  
 FT BINDING 43 48 ATP.  
 FT ACT\_SITE 136 136 BY SIMILARITY.  
 FT VARSPPLIC 1 12 MNPGLSRRNP -> MA (in isoform 2).  
 FT VARSPPLIC 352 372 Missing (in isoform 3).  
 FT MUTAGEN 48 48 K->E: LOSS OF KINASE ACTIVITY AND ABILITY TO ACTIVATE JNK FAMILY.  
 FT CONFLICT 392 392 N -> D (IN REF. 1; AAC15472).  
 SQ SEQUENCE 894 AA; 101315 MW; 6EH75BB34ESB733 CRC64;  
 Query Match 19.7%; Score 74; DB 1; Length 894;  
 Best Local Similarity 32.8%; Pred. No. 10;  
 Matches 21; Conservative 6; Mismatches 23; Indels 14; Gaps 3;  
 QY 1 PPFAPQVDSST-----GVHSSQVSGQAVTVSRPSLNAYNSLRSGLKRTPLSLKPDVPPKP 56  
 ID 432 PPLPLPKPKSLPIFQPMHSIEDENOG--TKRCP-----MSGSPAKPSQVPPRP 481  
 QY 57 STAP 56  
 ID 482 RIPP 485  
 RESULT 11  
 PCOL\_RAT  
 ID PCOL\_RAT STANDARD; PRT: 5085 AA.  
 AC Q9JKS5; Q9JL11;  
 DI 28-FEB-2003 (Rel. 41; Created)  
 DI 28-FEB-2003 (Rel. 41; Last sequence update)  
 DI 28-FEB-2003 (Rel. 41; Last annotation update)  
 DE Piccolo protein (Multidomain presynaptic cytomatrix protein).  
 GN PCLO.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 CX NCBI\_taxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH KABC1.

RX MEDLINE 20170257; PubMed-10707904;  
 RA Fester S.D., Chung W.J., Zhai R., Cases-Landhoff C., Voss B.,  
 RA Garner A.M., Kacmpf U., Kinder S., Gundelfinger E.D., Garner C.C.,  
 RT "Piccolo, a presynaptic zinc finger protein structurally related to  
 RT bassoon", a presynaptic zinc finger protein structurally related to  
 RL bassoon", a presynaptic zinc finger protein structurally related to  
 RL Neuron 25:203-214(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Fester S.D., Cases-Landhoff C., Gundelfinger E.D., Garner C.C.,  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP CALCIUM-BINDING ACTIVITY, AND MUTAGENESIS OF ASP-4668; ASP-4674;  
 RP VAL-4688; MET-4689; VAL-4693; SER-4693; GLN-4692; ASN-4693 AND  
 RP ALA-4694.  
 RX MEDLINE-21181819; PubMed-11285225;  
 RA Gerber S.H., Garcia J., Rizo J., Suchofsky J.B.,  
 RT "An unusual C(2)-domain in the active-zone protein piccolo:  
 RT implications for Ca(2+) regulation of neurotransmitter release",  
 RL EMBO J. 20:1605-1619(2001).  
 CC -!- FUNCTION: May act as a scaffolding protein involved in the  
 CC organization of synaptic active zones and in synaptic vesicle  
 CC trafficking (by similarity).  
 CC -!- SUBUNIT: Interacts with Rabact/Pral and profilin.  
 CC -!- SUBCELLULAR LOCATION: Concentrated at presynaptic side of synaptic  
 CC junctions.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-A: Alternative splicing; Named isoforms 1-2;  
 CC Name-1:  
 CC IsoId-Q9JXS6-1; Sequence-Displayed;  
 CC Name-2:  
 CC IsoId-Q9JXS6-2; Sequence-VSP\_004930; VSP\_003931;  
 CC -!- DOMAIN: C2 domain 1 is involved in binding calcium and  
 CC phospholipids. Calcium binds with low affinity but with high  
 CC specificity and induces a large conformational change.  
 CC -!- SIMILARITY: Contains 2 C2 domains.  
 CC -!- SIMILARITY: Contains 1 PDZ/CRK domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR ENBL: AF138789; AAF07822.2; -;  
 DR HSSP: AF227534; AAF63196.1; -;  
 DR GO: GO:0045202; C:synaptic junction; IDA.  
 DR GO: GO:0005509; F:calcium ion binding activity; IDA.  
 DR GO: GO:0005544; F:calcium-dependent phospholipid binding activity; IDA.  
 DR GO: GO:0005522; F:profilin binding activity; ISS.  
 DR GO: GO:0007010; P:cytoskeleton organization and biogenesis; ISS.  
 DR GO: GO:0016080; P:synaptic vesicle targeting; NAS.  
 DR InterPro: IPR003008; C2.  
 DR InterPro: IPR001478; PDZ.  
 DR Pfam: PF00168; C2; 2.  
 DR Pfam: PF00595; PDZ; 1.  
 DR SMART: SM00239; C2; 2.  
 DR SMART: SM00228; PDZ; 1.  
 DR PROSITE: PS00499; C2\_DOMAIN\_1; 1.  
 DR PROSITE: PS00004; C2\_DOMAIN\_2; 2.  
 DR PROSITE: PS0106; PDZ; 1.  
 KW Calcium/phospholipid-binding; Metal-binding; Zinc; Zinc-finger;  
 KW Repeat; Alternative splicing; 12 x 10 AA TANDEN APPROXIMATE REPEATS OF  
 FT DOMAIN 372 491  
 FT P-A-K-P-Q-P-Q-P-X.  
 FT ZN\_FING 523 547  
 FT ZN\_FING 1010 1033  
 FT DOMAIN 2351 2362  
 FT DOMAIN 4342 4536  
 FT DOMAIN 4653 4752  
 FT DOMAIN 4968 5059  
 FT C2 DOMAIN 1.  
 FT C2 DOMAIN 2.

FT VARSPLIC 4875 4880  
 FT VARSPLIC 4881 5085  
 FT MUTAGEN 4668 4668  
 FT MUTAGEN 4674 4674  
 FT MUTAGEN 4688 4688  
 FT MUTAGEN 4689 4689  
 FT MUTAGEN 4690 4691  
 FT MUTAGEN 4692 4693  
 FT MUTAGEN 4694 4694  
 FT SEQUENCE 5085 AA; 552702 MW; 5A1BH543201A7450 CRC64;  
 Query Match 19.5%; Score 73.5; DB 1; Length 5085;  
 Best Local Similarity 26.4%; Pred. No. 81;  
 Matches 19; Conservative 8; Mismatches 20; Indels 25; Gaps 3;  
 QY 1 PPTAFQVLSGVHSSQSPSSQAVTVSKOPS-NAYNSLGRSLKFTFSLKIUVPKPSFAP 60  
 ID 1 PPTAFQVLSGVHSSQSPSSQAVTVSKOPS-NAYNSLGRSLKFTFSLKIUVPKPSFAP 60  
 QY 51 LS-TSMKPNAC 71  
 ID 51 LS-TSMKPNAC 71  
 QY 2433 IVTTQKPTQTC 2479  
 ID 2433 IVTTQKPTQTC 2479  
 RESULT 12  
 V70K\_TYVW STANDARD; PRT: 628 AA.  
 AC P10357;  
 DT 01-MAR-1989 (rel. 10; Created)  
 DT 01-AUG-1992 (rel. 23; Last sequence update)  
 DT 01-AUG-1992 (rel. 23; Last annotation update)  
 DE 69 kDa protein.  
 CS Turnip yellow mosaic virus.  
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymovirus.  
 CX NCBI:taxid-12554;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-88289359; PubMed-3359388;  
 RA Morch M.D., Boyer J.C., Hachni A.,  
 RT "Overlapping open reading frames revealed by complete nucleotide  
 RT sequencing of turnip yellow mosaic virus genomic RNA.",  
 RL Nucleic Acids Res. 16:6157-6173(1988).  
 CC -!- FUNCTION: NOT KNOWN.  
 CC -!- SIMILARITY: TO 65 TO 70 KDA PROTEIN FROM OTHER TYMOVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR ENBL: X07441; CAA30321.1; ALT\_SEQ.  
 DR PIR: S01955; S01955.  
 DR InterPro: IPR004935; Tymo\_45-70kDa.  
 DR Pfam: PF03251; Tymo\_45kd\_70kd; 1.  
 DR SEQUENCE 628 AA; 69195 MW; 9B01CB5ADECEAC77 CRC64;  
 Query Match 19.3%; Score 72.5; DB 1; Length 628;  
 Best Local Similarity 29.6%; Pred. No. 9.8;



```

FT REPEAT      465      484      13.
FT REPEAT      485      503      14.
FT REPEAT      504      523      15.
FT REPEAT      524      543      16.
SQ SEQUENCE    558 AA: 512BC363dBA9F3C CRG64;
      Query Match      19.0%; Score 71.5; DB 1; Length 668;
      Best local Similarity 30.0%; Pred. No. 13;
      Matches 24; Conservative 11; Mismatches 20; Indels 25; Gaps 5;
OY 10 SIQVHSCQPSGQ-----AVTVSRQ-----PSLNAYNSLTFESGL-KETPSL 48
DB 127 SLEVLKQPCGSNANTENTIASATTPVKRKQSWPHPTFGYNASSSTGTRREPCV 146
OY 49 KPDVPPKPSAPSTSKPN 68
DB 187 HPPIPPKPS-SPVSS---PN 262

RESULT 15
PCL1_HUMAN
ID PCL1_HUMAN STANDARD; PRT: 1004 AA.
AC P78364; ORFWD3;
DI 28-FEB-2003 (Ref. 41; Created)
DI 28-FEB-2003 (Ref. 41; Last sequence update)
DI 15-SEP-2003 (Ref. 42; Last annotation update)
DE Polytomeotic-like protein 1 (Early development regulator protein 1)
DE (HPH1).
GN PCL1 OR EDRI OR PH1.
CS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae;
OC NCBI_TaxID=9606;
XP MEDLINE-9722024; PubMed=9211482;
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND INTERACTION WITH BM11.
RA Gunster M.J., Satlin D.P., Haver K.M., den Blaauwen J.,
de Bruijn D., Alkema M.J., van Loonhuizen M., van Griel R., Otte A.F.,
"Identification and characterization of interactions between the
vertebrate polycomb-group protein BM11 and human homologs of
polytomeotic."
RL Mol. Cell. Biol. 17:2826-2835(1997).
RN [2]
RP SEQUENCE OF 504-1004 FROM N.A.
RC TISSUE=Lung and Lymph.
RA Strausberg R.L., Feingold E.A., Grouse L.H., Berger J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.J., Wang J., Hsieh F.,
Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong I.,
Stapleton M., Soares M.B., Bonaldo M.P., Casavant J.L., Schetz T.E.,
Brownstein M.J., Udell I.B., Ishiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallaby S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guaratine F.H.,
Richards S., Worley K.J., Hale S., Garcia A.M., Gay J.J., Hulyk S.W.,
Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs B.A.,
Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
Whitling M., Madan A., Young A.C., Shevchenko Y., Beffard G.G.,
Blakesley R.W., Touchman C.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.J., Skalska J., Smalls R.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
"Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Involved in the regulation of Box gene expression and
segment specification during paraxial mesoderm and neural crest
development (By similarity).
CC -!- SUBUNIT: Homodimer. Interacts with BM11.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to licensedsib-sib.ch).
DR EMBL: U09277; AAC53169.1;
DR EMBL: BC002871; AAH02871.1; ALT_INIT.
DR EMBL: BC017748; AAH17748.1; ALT_INIT.
DR Genbank: F0303182; P031.
DR MIM: 602978;
DR GO: GO:0005634; Cytocleus; TAS.
DR GO: GO:0005515; Fibrinogen binding activity; TAS.
DR InterPro: IPR001663; SAM.
DR Pfam: PF00535; SAM.
DR SMART: SM00454; SAM.
DR PROSITE: PS5105; SAM_DOMAIN; 1.
KW Developmental protein; DNA-binding; Nuclear protein; Metal-binding;
KW Zinc-finger.
FT DOMAIN      425      441      POLY-GLN.
FT DOMAIN      940      1004      SAM.
FT 2N-FING     809      823      C4-TYPE (POTENTIAL).
FT CONFLICT    566      569      HL -> LK (IN REF. 1).
FT CONFLICT    613      613      S -> T (IN REF. 1).
FT CONFLICT    693      693      A -> T (IN REF. 1).
FT CONFLICT    752      752      V -> G (IN REF. 1).
FT CONFLICT    782      782      L -> F (IN REF. 1).
FT CONFLICT    972      972      L -> F (IN REF. 1).
SQ SEQUENCE    1004 AA: 105418.MW: 5E35765759904C4F CRG64;
      Query Match      19.0%; Score 71.5; DB 1; Length 1004;
      Best Local Similarity 35.8%; Pred. No. 21;
      Matches 24; Conservative 5; Mismatches 23; Indels 15; Gaps 3;
OY 2 PPAQGVPSIOVHSCQPSGQAVTVSRQPSLNAYNSLTFESGLKRTTP-SLKPTVPKPSFAP 60
DB 448 PDPVGVPTQVPTFSCSQDQACTLVQPELOS-----SPLSLPPPAKPK---P 493
OY 61 LSTSMKP 67
DB 494 LPQSKP 500

```

Search completed: September 30, 2003, 16:33:00  
Job time : 40 secs



GenCore version 5.1.6  
Copyright: (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw mode.

Run on: September 30, 2003, 16:19:53 : Search time 69 seconds  
(without alignments)  
269.272 Million cell updates/sec

Title: US-09-856-681-4

Perfect score: 376

Sequence: 1 PPPAPQVRVDSIQVHSSQPSGOAVTVSQPSLNATVRSGLKRTPSLKPDVPPKPSFAP 60

Scoring table: BL2SUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 254552604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 04

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:

- 1: sp\_archaea:
- 2: sp\_bacteria:
- 3: sp\_fungi:
- 4: sp\_human:
- 5: sp\_invertebrate:
- 6: sp\_mammal:
- 7: sp\_mnc:
- 8: sp\_organelle:
- 9: sp\_phage:
- 10: sp\_plant:
- 11: sp\_rodent:
- 12: sp\_virus:
- 13: sp\_vertebrate:
- 14: sp\_unclassified:
- 15: sp\_virus:
- 16: sp\_bacteria:
- 17: sp\_archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	376	100.0	507	4 Q96T04	Q96T04 homo sapien
2	376	100.0	562	4 Q96SY4	Q96SY4 homo sapien
3	376	100.0	562	4 Q8NC49	Q8NC49 homo sapien
4	376	100.0	574	4 Q96SM8	Q96SM8 homo sapien
5	376	100.0	699	4 Q96SM4	Q96SM4 homo sapien
6	376	100.0	1005	11 Q96Q71	Q96Q71 mus musculu
7	376	100.0	1030	4 Q96ZP5	Q96ZP5 homo sapien
8	376	100.0	1049	4 Q96ZHG	Q96ZHG homo sapien
9	163.5	43.5	416	6 Q95KAG	Q95KAG macaca fasc
10	163.5	43.5	451	4 Q9H9K4	Q9H9K4 homo sapien
11	163.5	43.5	454	4 Q9H9G5	Q9H9G5 homo sapien
12	163.5	43.5	998	4 Q8NFY6	Q8NFY6 homo sapien
13	163.5	43.5	1011	4 Q8NFY3	Q8NFY3 homo sapien
14	163.5	43.5	1017	4 Q8NFY5	Q8NFY5 homo sapien
15	163.5	43.5	1022	4 Q9P249	Q9P249 homo sapien
16	163.5	43.5	1073	4 Q8NFY4	Q8NFY4 homo sapien

#### ALIGNMENTS

RESULT :

Q96T04 PRELIMINARY: PRI: 507 AA.  
A: Q96T04:  
D: 01-DEC-2001 (TrEMBLrel. 19, Created)  
D: 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
D: 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein FLJ14533.  
DS Homo sapiens (Human).  
EC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
SC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Watanabe S., Hosoki T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Katanabe S., Kimura K., Marakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
RA Ninomiya K., Iwayanagi T.  
RA \*NEDO human cDNA sequencing project.\*  
RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AK027439; BAB51111.1;  
DR InterPro; IPR003659; Plexin-like.  
DR SMART; SM00423; PSI: 1.  
KW Hypothetical protein.  
SQ SEQUENCE 507 AA; 55464 MW; 8CC567B438C51B39 CRC64;

Query Match 100.0%; Score 376; DB 4; Length 507;  
Best Local Similarity 100.0%; Pred. No. 5.6e-33;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPAPQVRVDSIQVHSSQPSGOAVTVSQPSLNATVRSGLKRTPSLKPDVPPKPSFAP 60  
|||||  
2b 436 PPPAPQVRVDSIQVHSSQPSGOAVTVSQPSLNATVRSGLKRTPSLKPDVPPKPSFAP 495  
Qy 61 LSTSMKPNDACT 72  
|||||

DB 496 LSTSMKPNDACT 562

## RESULT 2

Q96SY4

ID Q96SY4 PRELIMINARY: PRT: 562 AA.  
AC Q96SY4 (TREMURel. 19, Created)  
DT 01-DEC-2001 (TREMURel. 19, Last sequence update)  
DE 01-MAR-2003 (TREMURel. 23, Last annotation update)  
DE Hypothetical protein FLJ14545.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagatsuma M., Hosoiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Naganari K., Masuho Y.,  
RA Niimiya K., Iwayanagi T.,  
RI "NEDO human cDNA sequencing project."  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK027471; BAH5136.1;  
DR InterPro: IPR003659; Plexin-like.  
DR SMART: SM00423; PS2.1;  
KW Hypothetical protein.  
SQ SEQUENCE 562 AA: 6133 MW: 6AB36957A7D5D78A CRC64:

Query Match 100.0% Score 376; DB 4; Length 562;  
Best Local Similarity 100.0%; Pred. No. 6.3e-33;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPAPGVDSIOVHSSQSGQAVTVSRQPSLNAYNSLRSGIKRTPSLKPDVPPKPSFAP 60  
DB 491 PPAPGVDSIOVHSSQSGQAVTVSRQPSLNAYNSLRSGIKRTPSLKPDVPPKPSFAP 550

QY 61 LSTSMKPNDACT 72

DB 551 LSTSMKPNDACT 562

## RESULT 3

Q96C49

ID Q96C49 PRELIMINARY: PRT: 562 AA.  
AC Q96C49 (TREMURel. 22, Created)  
DT 01-OCT-2002 (TREMURel. 22, Last sequence update)  
DE 01-OCT-2002 (TREMURel. 22, Last annotation update)  
DE Hypothetical protein FLJ93494.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,  
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kodaira H., Sudo H.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kodaira H., Sudo H.,  
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Kojima S., Naganari K.,  
RA Hattori A., Okumura K., Iwayanagi T., Niimiya K.,  
RI "NEDO human cDNA sequencing project."  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK074575; BAC11326.1;  
KW Hypothetical protein.

Query Match 100.0% Score 376; DB 4; Length 562;  
Best Local Similarity 100.0%; Pred. No. 6.3e-33;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPAPGVDSIOVHSSQSGQAVTVSRQPSLNAYNSLRSGIKRTPSLKPDVPPKPSFAP 60  
DB 491 PPAPGVDSIOVHSSQSGQAVTVSRQPSLNAYNSLRSGIKRTPSLKPDVPPKPSFAP 550

QY 61 LSTSMKPNDACT 72

DB 551 LSTSMKPNDACT 562

## RESULT 4

Q96SW4

ID Q96SW4 PRELIMINARY: PRT: 574 AA.  
AC Q96SW4 (TREMURel. 19, Created)  
DT 01-DEC-2001 (TREMURel. 19, Last sequence update)  
DE 01-MAR-2003 (TREMURel. 23, Last annotation update)  
DE Hypothetical protein FLJ14746.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagatsuma M., Hosoiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Naganari K., Masuho Y.,  
RA Niimiya K., Iwayanagi T.,  
RI "NEDO human cDNA sequencing project."  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK027654; BAH55265.1;  
DR InterPro: IPR003659; Plexin-like.  
DR SMART: SM00423; PS2.1;  
KW Hypothetical protein.

QY SEQUENCE 574 AA: 62822 MW: 0C79E01A4117A495 CRC64:

Query Match 100.0% Score 376; DB 4; Length 574;  
Best Local Similarity 100.0%; Pred. No. 6.5e-33;  
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPAPGVDSIOVHSSQSGQAVTVSRQPSLNAYNSLRSGIKRTPSLKPDVPPKPSFAP 60  
DB 503 PPAPGVDSIOVHSSQSGQAVTVSRQPSLNAYNSLRSGIKRTPSLKPDVPPKPSFAP 562

QY 61 LSTSMKPNDACT 72

DB 563 LSTSMKPNDACT 574

## RESULT 5

Q96SW4

ID Q96SW4 PRELIMINARY: PRT: 599 AA.  
AC Q96SW4 (TREMURel. 19, Created)  
DT 01-DEC-2001 (TREMURel. 19, Last sequence update)  
DE 01-MAR-2003 (TREMURel. 23, Last annotation update)  
DE Hypothetical protein FLJ14553.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagatsuma M., Hosoiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Naganari K., Masuho Y.,  
RA Niimiya K., Iwayanagi T.,

```

RT "NEDO human cDNA sequencing project."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL: AK027501; BAB55158.1; .
DR InterPro: IPR003659; Plexin-like.
DR InterPro: IPR001627; SEMA.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00423; PSI; 1.
KW Hypothetical protein.
SQ SEQUENCE 599 AA; 76723 MW; 2E5F11D59741394 CRC64;

Query Match 100.0%; Score 376; DB 4; Length 693;
Best Local Similarity 100.0%; Pred. No. 1.3e-42;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
DB 628 PPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 607
QY 61 LSTSMKPNDACT 72
DB 688 LSTSMKPNDACT 699

RESULT 6
Q9H2E6
ID Q9H2E6 PRELIMINARY; PRT: 1005 AA.
AC Q9H2E6
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 23, Last annotation update)
DE Axiol guidance signal SEMA6A.
GN SEMA6A.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20564339; PubMed=10993894;
RA Klosterman A., Lutz B., Gertler F., Behl C.;
RT "The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1/SEMA6A-1) bind to the Enabled/Vasodilator-stimulated Phosphoprotein-like Protein (EVL) via a novel carboxyl-terminal Zyxin-like domain."
RT like Protein (EVL) via a novel carboxyl-terminal Zyxin-like domain."
RL J. Biol. Chem. 275:39647-39653(2000).
DR EMBL: AF286666; AAC29434.1; .
DR InterPro: IPR003659; Plexin-like.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00630; Sema; 1.
SQ SEQUENCE 1305 AA; 111758 MW; 579699277F45B079D CRC64;

Query Match 100.0%; Score 376; DB 11; Length 1005;
Best Local Similarity 100.0%; Pred. No. 1.2e-42;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
DB 914 PPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 905
QY 61 LSTSMKPNDACT 72
DB 994 LSTSMKPNDACT 1005

RESULT 7
Q9H2E6
ID Q9H2E6 PRELIMINARY; PRT: 1030 AA.
AC Q9H2E6
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

```

```

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Semaphorin SEMA6A1.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20564339; PubMed=10993894;
RA Klosterman A., Lutz B., Gertler F., Behl C.;
RT "The orthologous human and murine semaphorin 6A-1 proteins (SEMA6A-1/SEMA6A-1) bind to the Enabled/Vasodilator-stimulated Phosphoprotein-like Protein (EVL) via a novel carboxyl-terminal Zyxin-like domain."
RL J. Biol. Chem. 275:39647-39653(2000).
DR EMBL: AF279656; AAC29378.1; .
DR InterPro: IPR003659; Plexin-like.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00630; Sema; 1.
SQ SEQUENCE 1030 AA; 114368 MW; A57B79C10AFC4B34 CRC64;

Query Match 100.0%; Score 376; DB 4; Length 1030;
Best Local Similarity 100.0%; Pred. No. 1.3e-42;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60
DB 959 PPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 1618
QY 61 LSTSMKPNDACT 72
DB 1019 LSTSMKPNDACT 1030

RESULT 8
Q9H2E6
ID Q9H2E6 PRELIMINARY; PRT: 1049 AA.
AC Q9H2E6
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein KIAA1368 (Fragment).
GN KIAA1368
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20181126; PubMed=10718196;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Chazawa M., Chazawa M.;
RT "Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro."
RL DNA Res. 7:65-73(2000).
DR EMBL: AAC37789; BAA92606.1; .
DR Genes; HGNC:10738; SEMA6A.
DR InterPro: IPR003659; Plexin-like.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00423; PSI; 1.
DR SMART: SM00630; Sema; 1.
KW Hypothetical protein.
FT NONTER
SQ SEQUENCE 1045 AA; 114511 MW; 7781D20ACC7A9AEA CRC64;

Query Match 100.0%; Score 376; DB 4; Length 1049;
Best Local Similarity 100.0%; Pred. No. 1.3e-42;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPAPQVDSIQVHSSQPSGQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAP 60

```



```

DR SMART: SMO0630; Sema: 1.
SQ SEQUENCE 998 AA: 111710 MW: 3F46D6872EB9D5544 CRC64:

Query Match 43.5%; Score 163.5; DB 4; Length 998;
Best Local Similarity 50.7%; Pred. No. 2.4e-09;
Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

QY 1 PPPAPQVRLSTGVHSSQPSGOAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDVPP 54
DB 927 PPTGAKVDYIQ-----GTPVSVHLQPSLSKROSSYTSNGTLPRTCGLKRTPSLKPDVPP 979
QY 55 KPSFAPLSTSMKP 67
DB 993 KPSEVPQTSPVRP 1005

RESULT 13
Q8NFY3 PRELIMINARY: PRT: 1017 AA.
AC Q8NFY3
DT 01-OCT-2002 (TrEMBLrel: 22, Created)
DT 01-OCT-2002 (TrEMBLrel: 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel: 23, Last annotation update)
DE Semaphorin 6D isoform 1.
GN SEMA6D.
CS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ou X., Wei H., Zhai Y., Yu Y., Tang F., He F.
RT "Homo sapiens semaphorin 6D isoform 1 (SEMA6D.1): cDNA, complete cds."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBS databases.
DR EMBL: AF389430; AA069453.1;
DR InterPro: IPR001627; Sema.
DR Pfam: PF01403; Sema: 1.
DR SMART: SMO0630; Sema: 1.
SQ SEQUENCE 1011 AA: 113269 MW: 9D6B8B3631941589 CRC64:

Query Match 43.5%; Score 163.5; DB 4; Length 1011;
Best Local Similarity 50.7%; Pred. No. 2.4e-09;
Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

QY 1 PPPAPQVRLSTGVHSSQPSGOAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDVPP 54
DB 940 PPTGAKVDYIQ-----GTPVSVHLQPSLSKROSSYTSNGTLPRTCGLKRTPSLKPDVPP 992
QY 55 KPSFAPLSTSMKP 67
DB 993 KPSEVPQTSPVRP 1005

RESULT 14
Q8NFY5 PRELIMINARY: PRT: 1017 AA.
AC Q8NFY5
DT 01-OCT-2002 (TrEMBLrel: 22, Created)
DT 01-OCT-2002 (TrEMBLrel: 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel: 23, Last annotation update)
DE Semaphorin 6D isoform 3.
GN SEMA6D.
CS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ou X., Wei H., Zhai Y., Yu Y., Tang F., He F.
RT "Homo sapiens semaphorin 6D isoform 3 (SEMA6D.3): cDNA, complete cds."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBS databases.

```

```

DR EMBL: AF389428; AA069451.1;
DR InterPro: IPR001627; Sema.
DR Pfam: PF01403; Sema: 1.
DR SMART: SMO0630; Sema: 1.
SQ SEQUENCE 1017 AA: 113736 MW: 4D639CEBADD9F2A0 CRC64:

Query Match 43.5%; Score 163.5; DB 4; Length 1017;
Best Local Similarity 50.7%; Pred. No. 2.4e-09;
Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

QY 1 PPPAPQVRLSTGVHSSQPSGOAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDVPP 54
DB 946 PPTGAKVDYIQ-----GTPVSVHLQPSLSKROSSYTSNGTLPRTCGLKRTPSLKPDVPP 998
QY 55 KPSFAPLSTSMKP 67
DB 999 KPSEVPQTSPVRP 1011

RESULT 15
Q8NFY5 PRELIMINARY: PRT: 1022 AA.
AC Q8NFY5
DT 01-OCT-2000 (TrEMBLrel: 15, Created)
DT 01-OCT-2000 (TrEMBLrel: 18, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel: 24, Last annotation update)
DE Hypothetical protein K1A1479 (Fragment).
GN K1A1479.
CS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-2027192; PubMed=8819331;
RA Nagase T., Kikuno R., Ishikawa K., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human
RT genes XVII. The complete sequences of 100 new cDNA clones from brain
RT which code for large proteins in vitro."
RL DNA Res. 7:143-150(2000).
DR EMBL: AR040912; BA055003.2;
DR InterPro: IPR003659; Plexin-like.
DR Pfam: PF01403; Sema.
DR SMART: SMO0424; Sema: 1.
DR SMART: SMO0630; Sema: 1.
KW Hypothetical protein.
ST NON-TER
SQ SEQUENCE 1022 AA: 114372 MW: BE4FRD5EA02C59C4 CRC64:

Query Match 43.5%; Score 163.5; DB 4; Length 1022;
Best Local Similarity 50.7%; Pred. No. 2.4e-09;
Matches 37; Conservative 8; Mismatches 15; Indels 13; Gaps 2;

QY 1 PPPAPQVRLSTGVHSSQPSGOAVTVSRQPSLNAYNSLT-----RSGLKRTPSLKPDVPP 54
DB 951 PPTGAKVDYIQ-----GTPVSVHLQPSLSKROSSYTSNGTLPRTCGLKRTPSLKPDVPP 1003
QY 55 KPSFAPLSTSMKP 67
DB 1004 KPSEVPQTSPVRP 1016

Search completed: September 30, 2003, 16:34:26
Job time : 75 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Computer, Ltd.

CM protein - protein search, using sw model

Run on: September 30, 2003, 16:31:08 : Search time 42 seconds  
(without alignments)  
72.533 Million full updates/sec

Title: us-09-856-681-4

Perfect score: 376

Sequence: 1 PPAAPQVDSIQVHSSQPSG.....PPKPSAPLSTSMKPNDACT 72

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 4231858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2268000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

#### Database :

Issued\_Patents\_AA:

1: /cqn2\_5/ptdata/1/iaa/5A\_COMB.pep.\*  
2: /cqn2\_5/ptdata/1/iaa/5B\_COMB.pep.\*  
3: /cqn2\_5/ptdata/1/iaa/5A\_COMB.pep.\*  
4: /cqn2\_5/ptdata/1/iaa/5B\_COMB.pep.\*  
5: /cqn2\_5/ptdata/1/iaa/PCBUS\_COMB.pep.\*  
6: /cqn2\_5/ptdata/1/iaa/backlist.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	72	19.1	480	3	US-09-189-035-5
2	72	19.1	480	3	US-09-382-086-5
3	71.5	19.0	1004	3	US-08-916-352-2
4	69.5	18.5	1125	4	US-08-913-783A-152
5	69.5	18.5	1610	4	US-09-513-783A-22
6	69	18.4	400	4	US-09-252-991A-32313
7	68.5	18.2	525	4	US-09-196-270-6
8	68	18.1	1149	3	US-08-560-005-5
9	68	18.1	1149	3	US-08-416-540-5
10	68	18.1	1149	4	US-09-969-528-5
11	68	18.1	1253	1	US-08-252-956B-12
12	68	18.1	1261	1	US-08-252-956B-18
13	67	17.8	169	4	US-09-252-991A-32563
14	67	17.8	2035	1	US-08-046-585-5
15	67	17.8	2035	1	US-08-393-703-5
16	67	17.8	2035	5	PCT-US93-11721-5
17	65.5	17.4	167	4	US-09-252-991A-32720
18	65.5	17.4	366	4	US-09-252-991A-32385
19	65.5	17.4	1142	2	US-08-993-116-7
20	65.5	17.4	1142	3	US-08-845-528C-7
21	65.5	17.4	1142	3	US-09-061-709-2
22	65.5	17.4	1142	4	US-09-566-2815-7
23	65.5	17.4	1242	4	US-09-499-651-2
24	65.5	17.4	1297	3	US-09-540-245A-17
25	65	17.3	70	4	US-09-006-428A-15
26	65	17.3	143	4	US-09-252-991A-22566
27	65	17.3	351	3	US-08-466-465-6

28	64.5	17.2	433	3	US-09-046-158A-2	Sequence 2, Appli
29	64.5	17.2	897	1	US-07-960-389-2	Sequence 2, Appli
30	64.5	17.2	1068	4	US-09-233-857-13	Sequence 13, Appli
31	64.5	17.2	1099	4	US-09-442-100-2	Sequence 2, Appli
32	64.5	17.2	1099	4	US-08-939-106-2	Sequence 2, Appli
33	64.5	17.2	3963	3	US-08-061-376-5	Sequence 5, Appli
34	64	17.0	641	4	US-09-071-035-456	Sequence 456, App
35	64	17.0	1313	4	US-09-071-035-450	Sequence 450, App
36	64	17.0	1313	4	US-09-071-035-454	Sequence 454, App
37	63.5	16.9	280	4	US-09-252-991A-20783	Sequence 20783, A
38	63.5	16.9	302	4	US-09-252-991A-21231	Sequence 21231, A
39	63.5	16.9	479	4	US-09-252-991A-32594	Sequence 32594, A
40	63	16.8	816	2	US-08-785-310A-8	Sequence 8, Appli
41	63	16.8	816	2	US-08-816-693A-53	Sequence 53, Appli
42	63	16.8	816	3	US-08-885-291-53	Sequence 53, Appli
43	63	16.8	816	3	US-09-496-672-53	Sequence 53, Appli
44	62.5	16.5	846	4	US-09-858-664A-3	Sequence 3, Appli
45	62	16.5	304	4	US-09-493-343B-22	Sequence 22, Appli

#### ALIGNMENTS

RESULT 1  
US-09-189-035-5  
: Sequence 5, Application US/09189035  
: Patent No. 6020165  
: GENERAL INFORMATION:  
: APPLICANT: Yue, Henry  
: APPLICANT: Corley, Neil C.  
: APPLICANT: Guegler, Karl J.  
: APPLICANT: Baughn, Mariah R.  
: TITLE OF INVENTION: CYCLIC SIGNAL REGULATORS  
: FILE REFERENCE: PP-0638 US  
: CURRENT APPLICATION NUMBER: US/09/189,035  
: CURRENT FILING DATE: 1998 11-10  
: NUMBER OF SEQ ID NOS: 6  
: SOFTWARE: PERL Program  
: SEQ ID NO 5  
: LENGTH: 460  
: TYPE: PRT  
: ORGANISM: Homo sapiens  
: FEATURE: -  
: OTHER INFORMATION: 32245671  
US-09-189-035-5

Query Match 19.1% Score 72; DB 3; Length 480;  
Best Local Similarity 31.2%; Pred. No. 4.8;  
Matches 24; Conservative 13; Mismatches 26; Indels 14; Gaps 4;  
QY 1 PPAAPQVDSIQVHSSQPSGQAVTSRPSLNAYNSLTRS-----GLKRTPSL-----KP 50  
DB 264 PTPSPPIIG--PAPCSAPSSQYGTMTIQ--ISRHNSTTSSTSSGYPRTFTSVTAQFSAP 319  
QY 51 DVPPKPSAPLSTSMKP 67  
DB 320 HVNGSPYQNSLSIAP 336

RESULT 2  
US-09-382-086-5  
: Sequence 5, Application US/09382086  
: Patent No. 620106  
: GENERAL INFORMATION:  
: APPLICANT: Yue, Henry  
: APPLICANT: Corley, Neil C.  
: APPLICANT: Guegler, Karl J.  
: APPLICANT: Baughn, Mariah R.  
: TITLE OF INVENTION: CYTOKINE SIGNAL REGULATORS  
: FILE REFERENCE: PP-0638 US  
: CURRENT APPLICATION NUMBER: US/09/382,086  
: CURRENT FILING DATE: 1999-08-24  
: EARLIER APPLICATION NUMBER: 09/189,035

```

: EARLIER FILING DATE: 1998-11-10
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PERL Program
: SEQ ID NO 5
: LENGTH: 480
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURES:
: OTHER INFORMATION: g2245671
US-09-382-086-5

Query Match      19.1%; Score 71.5; DB 3; Length 1004;
Best Local Similarity 31.2%; Pred. No. 14;
Matches 24; Conservative 5; Mismatches 23; Indels 15; Gaps 3;

QY 1 PPAPQVDSIQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAPL 60
DB 204 PTPSPPTIG--PAPGSAFGSGVGMTRQ--ISRHNSTTS--SSGGRYKRTPSVTAQPSAQ 60
QY 51 DVEPKPSFAPLSTSKP 67
DB 320 RVNKGPLSQNSISLAP 136

RESULT 3
US-08-916-352-2
: Sequence 2, Application US/08916352
: Patent No. 6166191
: GENERAL INFORMATION:
: APPLICANT: CHIRON CORPORATION
: TITLE OF INVENTION: HUMAN FOLYHOMEOGIC 1 (hph1) ACTS AS A
: TITLE OF INVENTION: TUMOR SUPPRESSOR
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CHIRON CORPORATION
: STREET: 4560 HORTON STREET
: CITY: EMERYVILLE
: STATE: CA
: COUNTRY: USA
: ZIP: 94608
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/06/316,352
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: FOTTER, JANE
: REGISTRATION NUMBER: 33,332
: REFERENCE/DOCKET NUMBER: 1355.
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 510-923-2707
: TELEFAX: 510-655-3542
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1004 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-916-352-2

Query Match      19.0%; Score 71.5; DB 3; Length 1004;
Best Local Similarity 35.8%; Pred. No. 14;
Matches 24; Conservative 5; Mismatches 23; Indels 15; Gaps 3;

QY 2 PPAPQVDSIQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAPL 60
DB 448 POPQVPTQGVPPSGSQDAGTIVQPMLOS-----SPLSIFPDAAKPK---P 493
```

```

QY 61 LSTSKAP 67
DB 494 IPIQSKP 500

RESULT 4
US-09-513-783A-152
: Sequence 22, Application US/09513783A
: Patent No. 6416959
: GENERAL INFORMATION:
: APPLICANT: Giuliano, Kenneth A.
: TITLE OF INVENTION: A System for Cell Based Screening
: FILE REFERENCE: 97-022-11
: CURRENT APPLICATION NUMBER: US/09/513,783A
: CURRENT FILING DATE: 2000-02-25
: NUMBER OF SEQ ID NOS: 180
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 152
: LENGTH: 1125
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-513-783A-152

Query Match      18.5%; Score 69.5; DB 4; Length 1125;
Best Local Similarity 28.6%; Pred. No. 26;
Matches 18; Conservative 13; Mismatches 31; Indels 1; Gaps 1;

QY 3 PAP-QRVDISQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAPL 61
DB 616 PPSPLENLEOKETPGCSQPSQVSRQEQEAKAAGVTGNDITTPNKEPPSPKAKPL 675
QY 62 STS 64
DB 676 ATT 678

RESULT 5
US-09-513-783A-22
: Sequence 22, Application US/09513783A
: Patent No. 6416959
: GENERAL INFORMATION:
: APPLICANT: Giuliano, Kenneth A.
: TITLE OF INVENTION: A System for Cell Based Screening
: FILE REFERENCE: 97-022-11
: CURRENT APPLICATION NUMBER: US/09/513,783A
: CURRENT FILING DATE: 2000-02-25
: NUMBER OF SEQ ID NOS: 180
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 22
: LENGTH: 1610
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURES:
: OTHER INFORMATION: Description of Artificial Sequence:
: OTHER INFORMATION: EYFD-DEVD-KAP4-EBFP construct
US-09-513-783A-22

Query Match      18.5%; Score 69.5; DB 4; Length 1610;
Best Local Similarity 28.6%; Pred. No. 41;
Matches 18; Conservative 13; Mismatches 31; Indels 1; Gaps 1;

QY 3 PAP-QRVDISQVHSSQPSQAVTVSRQPSLNAYNSLTRSGLKRTPSLKPDVPPKPSFAPL 61
DB 652 PPSPLENLEOKETPGCSQPSQVSRQEQEAKAAGVTGNDITTPNKEPPSPKAKPL 921
QY 62 STS 64
DB 922 ATT 924

RESULT 6
```

```
US-09-252-931A-32313
: Sequence 32313, Application US/09252991A
: Patent No. 6551755
: GENERAL INFORMATION:
: APPLICANT: Marc J. R-benfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.135
: CURRENT APPLICATION NUMBER: US/09/252.991A
: PRIOR FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/374,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/394,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 32313
: LENGTH: 400
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32313

Query Match: 18.43; Score 69; DB 4; Length 400;
Best Local Similarity: 30.19; Pred. No. 8.2;
Matches: 25; Conservative 9; Mismatches 20; Indels 24; Gaps 5;

QY 2 PPAPQV-----DSIQVHSSQPSGAVTVSKPSSINAYNSLTRSGIKRTTP 47
DB 115 PPATPRRLRRGGRAGATGCGASAAACAP-SCVAIVGAP-----GALFTAG--RAH 145

QY 48 LKPPVPPKPSFAPLSTSMKINDA 70
DB 166 NRPOPPATSESPMS---KTPA 185

RESULT 7
US-09-196-270-6
: Sequence 6, Application US/09196270
: Patent No. 6506316
: GENERAL INFORMATION:
: APPLICANT: Hecht, Peter
: APPLICANT: Madden, Kevin
: APPLICANT: Fink, Gerald
: TITLE OF INVENTION: CHIMERIC PRE-ACTIVATED TRANSCRIPTION
: TITLE OF INVENTION: FACTORS
: FILE REFERENCE: 50078/004002
: CURRENT APPLICATION NUMBER: US/09/196,270
: CURRENT FILING DATE: 1998-11-19
: EARLIER APPLICATION NUMBER: 60/066,129
: EARLIER FILING DATE: 1997-11-19
: EARLIER APPLICATION NUMBER: 60/066,308
: EARLIER FILING DATE: 1997-11-21
: EARLIER APPLICATION NUMBER: 60/066,462
: EARLIER FILING DATE: 1997-11-24
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 6
: LENGTH: 625
: TYPE: PRT
: ORGANISM: Saccharomyces cerevisiae
US-09-196-270-6

Query Match: 16.28; Score 68.5; DB 4; Length 625;
Best Local Similarity: 25.84; Pred. No. 16;
Matches: 31; Conservative 13; Mismatches 49; Indels 49; Gaps 6;

QY 1 PPAPQVQVDSQVHSS-----QPSQGVV-----VSQV-24
DB 478 PPNAPS-YOQVCGSSSATANTATYVVRVLAFTGTSITELPPVHSHVAGVTVSGS 476

QY 29 -----PSENAYSLTRSGIKRTTPSLKPPVPPKPSFAPLS--TSMKPN-----GAVT 72
DB 477 QVAMPHYISVRAATSYSSSSCSILPPVQSKIPVPSRRINAKOTS,KKPWFPSINUKSCL 546
```

```
RESULT 8
US-08-560-005-5
: Sequence 5, Application US/08560005
: Patent No. 6001354
: GENERAL INFORMATION:
: APPLICANT: Pot, David A.
: APPLICANT: Williams, Lewis T.
: APPLICANT: Jefferson, Anne Bennett
: APPLICANT: Majerus, Philip W.
: TITLE OF INVENTION: No. 6001354el Grb2 Associating Protein and Nucleic
: TITLE OF INVENTION: Acids Encoding Therefor
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew
: STREET: One Market Plaza, Steuart Tower, Suite 2000
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94105
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA: US/08/560,005
: APPLICATION NUMBER: US/08/560,005
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Dow, Karen H.
: REGISTRATION NUMBER: 25,664
: REFERENCE/DOCKET NUMBER: 2307K-0624000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-326-2400
: TELEFAX: 415-326-2422
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 144 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FEATURE:
: NAME/KEY: Region
: LOCATION: 1..144
: OTHER INFORMATION: /note="51c"
US-08-560-005-5

Query Match: 18.18; Score 68; DB 3; Length 1149;
Best Local Similarity: 29.94; Pred. No. 40;
Matches: 23; Conservative 12; Mismatches 16; Indels 5; Gaps 3;

QY 2 PPAPQV---DSIQVHSSQPSG---QAVTVSKQPSLNAYNSLTRSGIKRTTP-SLKPPVPPK 55
DB 380 PPAPRAAPRSEPTPRKPKGAPGVAAPPKPSFNFPAYVYLVGVPHQLPPPPPS 939

QY 56 PSFAPLSTSMKPNDACT 72
DB 540 PARAPVPSATKNKVAIT 956

RESULT 9
US-09-418-540-5
: Sequence 5, Application US/09418540
: Patent No. 6296648
: GENERAL INFORMATION:
: APPLICANT: Pot, David A.
: APPLICANT: Williams, Lewis T.
: APPLICANT: Jefferson, Anne Bennett
: APPLICANT: Majerus, Philip W.
: TITLE OF INVENTION: No. 6296648el Grb2 Associating Protein and Nucleic
: TITLE OF INVENTION: Acids Encoding Therefor
: NUMBER OF SEQUENCES: 10
```



1 CORRESPONDENCE ADDRESS:  
2 ADDRESSEE: Townsend and Townsend and Crew  
3 STREET: One Market Plaza, Stewart Tower, Suite 2000  
4 CITY: San Francisco  
5 STATE: California  
6 COUNTRY: USA  
7 ZIP: 94105

8 COMPUTER READABLE FORM:  
9 MEDIUM TYPE: Floppy disk  
10 COMPUTER: IBM PC compatible  
11 OPERATING SYSTEM: PC-DOS/MS-DOS  
12 SOFTWARE: Patent: Release #1.0, Version #1.25

13 CURRENT APPLICATION DATA:  
14 APPLICATION NUMBER: US/09/418-540  
15 FILING DATE: 14-OCT-1995

16 CLASSIFICATION:

17 PRIOR APPLICATION DATA:

18 APPLICATION NUMBER: US 08/560,005

19 FILING DATE: 17-NOV-1995

20 ATTORNEY/AGENT INFORMATION:

21 NAME: Dow, Karen B.

22 REGISTRATION NUMBER: 29,684

23 REFERENCE/DOCKET NUMBER: 2307K-0624000

24 TELECOMMUNICATION INFORMATION:

25 TELEPHONE: 415-326-2400

26 TELEFAX: 415-326-2422

27 INFORMATION FOR SEQ ID NO: 5:

28 SEQUENCE CHARACTERISTICS:

29 LENGTH: 1149 amino acids

30 TYPE: amino acid

31 TOPOLOGY: linear

32 MOLECULE TYPE: protein

33 FEATURE:

34 NAME/KEY: Region

35 LOCATION: 1..1149

36 OTHER INFORMATION: /note= "51c"

37 US-09-418-540-5

38 Query Match 18.1%; Score 68; DB 3; Length 1149;

39 Best Local Similarity 29.9%; Pred. No. 40;

40 Matches 23; Conservative 12; Mismatches 36; Indels 6; Gaps 3;

41 QY 2 PPAQKV--DSIQVHSSCPSG--QAVTVSRPSLNAYNSITRSGLKRP-SLKPQVPPK 55

42 Db 880 PPAPRAAPREPLTPRLKPEGAPEPGVAAPPKNSFNPNPAYVVEGVPHLLPEPPS 919

43 QY 56 PSFAPLSTSMKPNDACT 72

44 Db 340 PARAPVPSATKNKVAIT 956

45 RESULT 10

46 US-09-969-528-5

47 Sequence 5, Application US/09/969528

48 Patent No. 5472,97

49 GENERAL INFORMATION:

50 APPLICANT: Pot, David A.

51 Williams, Lewis T.

52 Jefferson, Anne Bennett

53 Majerus, Philip W.

54 TITLE OF INVENTION: No. 6472,97el 3rd Association Protein and Nucleic

55 Acid Encoding Thereof

56 NUMBER OF SEQUENCES: 10

57 CORRESPONDENCE ADDRESS:

58 ADDRESSEE: Townsend and Townsend and Crew

59 STREET: One Market Plaza, Stewart Tower, Suite 2000

60 CITY: San Francisco

61 STATE: California

62 COUNTRY: USA

63 ZIP: 94105

64 COMPUTER READABLE FORM:

65 MEDIUM TYPE: Floppy disk

66 COMPUTER: IBM PC compatible

67 OPERATING SYSTEM: PC-DOS/MS-DOS  
68 SOFTWARE: Patent: Release #1.0, Version #1.25

69 CURRENT APPLICATION DATA:

70 APPLICATION NUMBER: US/09/969,528

71 FILING DATE: 01-Oct-2001

72 CLASSIFICATION: <UNKNOWN>

73 PRIOR APPLICATION DATA:

74 APPLICATION NUMBER: 08/560,005

75 FILING DATE: <UNKNOWN>

76 ATTORNEY/AGENT INFORMATION:

77 NAME: Dow, Karen B.

78 REGISTRATION NUMBER: 29,684

79 REFERENCE/DOCKET NUMBER: 2307K-0624000

80 TELECOMMUNICATION INFORMATION:

81 TELEPHONE: 415-326-2400

82 TELEFAX: 415-326-2422

83 INFORMATION FOR SEQ ID NO: 5:

84 SEQUENCE CHARACTERISTICS:

85 LENGTH: 1149 amino acids

86 TYPE: amino acid

87 TOPOLOGY: linear

88 MOLECULE TYPE: protein

89 FEATURE:

90 NAME/KEY: Region

91 LOCATION: 1..1149

92 OTHER INFORMATION: /note= "51c"

93 SEQUENCE DESCRIPTION: SEQ ID NO: 5:

94 US-09-969-528-5

95 Query Match 18.1%; Score 68; DB 4; Length 1149;

96 Best Local Similarity 29.9%; Pred. No. 40;

97 Matches 23; Conservative 12; Mismatches 36; Indels 6; Gaps 3;

98 QY 2 PPAQKV--DSIQVHSSCPSG--QAVTVSRPSLNAYNSITRSGLKRP-SLKPQVPPK 55

99 Db 880 PPAPRAAPREPLTPRLKPEGAPEPGVAAPPKNSFNPNPAYVVEGVPHLLPEPPS 939

100 QY 56 PSFAPLSTSMKPNDACT 72

101 Db 340 PARAPVPSATKNKVAIT 956

102 RESULT 11

103 US-08-252-966B-12

104 Sequence 12, Application US/08252966B

105 Patent No. 5624818

106 GENERAL INFORMATION:

107 APPLICANT: Eisenman, Robert N.

108 Applicant: Haffin, Peter J.

109 APPLICANT: Aver, Donald E.

110 TITLE OF INVENTION: Regulatory Proteins that Dimerize with

111 TITLE OF INVENTION: Mad or Max

112 NUMBER OF SEQUENCES: 19

113 CORRESPONDENCE ADDRESS:

114 ADDRESSES: Christensen, O'Connor, Johnson, and Kindness PLLC

115 STREET: 1420 Fifth Ave., Suite 2800

116 CITY: Seattle

117 STATE: Washington

118 COUNTRY: USA

119 ZIP: 98101-2347

120 COMPUTER READABLE FORM:

121 MEDIUM TYPE: Floppy disk

122 COMPUTER: IBM PC compatible

123 OPERATING SYSTEM: PC-DOS/MS-DOS

124 SOFTWARE: Patent: Release #1.0, Version #1.25

125 CURRENT APPLICATION DATA:

126 APPLICATION NUMBER: US/08/252,966B

127 FILING DATE: 01-JUN-1994

128 CLASSIFICATION: 435

129 ATTORNEY/AGENT INFORMATION:

130 NAME: Shelton, Dennis K.

131 REGISTRATION NUMBER: 26,997

132 REFERENCE/DOCKET NUMBER: FHCRI7694

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 682-8100
: TELEFAX: (206) 224-0779
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1253 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: DESCRIPTION: translation of msina cDNA; see Figure 23
: HYPOTHETICAL: YES
: ORIGINAL SOURCE:
: ORGANISM: Mus musculus
: US-08-252-966B-12

Query Match 18.1% Score 68; DB 1; Length 1261;
Best Local Similarity 30.7%; Pred. No. 44;
Matches 23; Conservative 7; Mismatches 35; Indels 10; Gaps 3;

QY 1 PPAPQVDSIQVHSSQPSQAVTVSKQPS--LNAYNSLIRSLKRTISLKPDVPPKPSF 58
DB 238 PQPPFQ-----HPSQSSQASPTPAQAPQPTAAKYSKPSQLQAHTPASQQTPLPPY 290
QY 59 A-PLSTSMKPNDACT 72
DB 291 ASRSPVPVQPHPTVT 305

RESULT 13
US-09-252-991A-30563
: Sequence 30563, Application US/09252991A
: Patent No. 6351795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubinfeld et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252.991A
: PRIOR FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-15
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 34142
: SEQ ID NO 30563
: LENGTH: 169
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
: US-09-252-991A-30563

Query Match 17.8% Score 67; DB 4; Length 169;
Best Local Similarity 27.4%; Pred. No. 4.6;
Matches 23; Conservative 14; Mismatches 29; Indels 18; Gaps 3;

QY 3 PAPQVDSIQVHSSQPSQAVTVSKQPSLNAYNSL-----IRSLKRTPSLK---- 49
DB 2 PQSRATISIPSSSSSPAKIACSGCKKNPNRAGSLTPAPSPCCSRISQSRSTRPTCG 61
QY 50 --PQVPPKPSAPLSTSMKPNDACT 71
DB 62 WTPTVPVPA---ASTSIRPTTRK 82

RESULT 14
US-08-046-585-5
: Sequence 5, Application US/08046585
: Patent No. 5453362
: GENERAL INFORMATION:
: APPLICANT: Lemarco, Kelly
: APPLICANT: Wilson, Angus
: APPLICANT: Heitz, Winship
: TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:
: TITLE OF INVENTION: HOST CELL FACTOR
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
: STREET: 4 Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: CA
: COUNTRY: USA
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 682-8100
: TELEFAX: (206) 224-0779
: INFORMATION FOR SEQ ID NO: 12:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1253 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: DESCRIPTION: translation of msina cDNA; see Figure 23
: HYPOTHETICAL: YES
: ORIGINAL SOURCE:
: ORGANISM: Mus musculus
: US-08-252-966B-12

Query Match 18.1% Score 68; DB 1; Length 1253;
Best Local Similarity 30.7%; Pred. No. 44;
Matches 23; Conservative 7; Mismatches 35; Indels 10; Gaps 3;

QY 1 PPAPQVDSIQVHSSQPSQAVTVSKQPS--LNAYNSLIRSLKRTISLKPDVPPKPSF 58
DB 238 PQPPFQ-----HPSQSSQASPTPAQAPQPTAAKYSKPSQLQAHTPASQQTPLPPY 290
QY 59 A-PLSTSMKPNDACT 72
DB 291 ASRSPVPVQPHPTVT 305

RESULT 12
US-08-252-966B-14
: Sequence 16, Application US/08252966B
: Patent No. 5624816
: GENERAL INFORMATION:
: APPLICANT: Eisenman, Robert N.
: APPLICANT: Hurlin, Peter J.
: APPLICANT: Ayer, Donald F.
: TITLE OF INVENTION: Regulatory Proteins that Dimerize with
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Christensen, O'Connor, Johnson, and Kindness PLLC
: STREET: 1420 Fifth Ave., Suite 2800
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98101-2347
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/03/252.966B
: FILING DATE: 01-JUN-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Shelton, Dennis K.
: REGISTRATION NUMBER: 26,397.
: REFERENCE/DOCKET NUMBER: FHCRI7694
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 682-8100
: TELEFAX: (206) 224-0779
: INFORMATION FOR SEQ ID NO: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1261 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: DESCRIPTION: translation of msina9 cDNA; see Figure 29A, B, C, D
: HYPOTHETICAL: YES
: ORIGINAL SOURCE:
```

SOFTWARE: PatentIn Release #1.0, Version #1.25

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/04/046,585

FILING DATE: 12-APR-1994

CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A.

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-57503-1/RAG

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 272299

INFORMATION FOR SEQ ID NO: 5

SEQUENCE CHARACTERISTICS:

LENGTH: 2035 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-046-585-5

## Query Match

Best Local Similarity 32.8%; Score 67; DB 1; Length 2035;

Matches 22; Conservative 10; Mismatches 27; Indels 8; Gaps 3;

```

QY      2 PPAPQVDSIQVHSSQPSG-----QAVTVSRQPSLNAYNSLTR--SGLKRTPSLKPDVPP 54
Db      1741 PSTVALLPSTATSLAPSNITVAPOPVVVASPAKIOAAALTEVANGIE-SLGVKPDIPP 1799

QY      55 KPSFAPL 61
Db      1800 PPSKAPM 1806

```

## RESULT 15

US-08-393-703-5

Sequence 5, Application US/08393703

Patent No. 5585239

## GENERAL INFORMATION:

APPLICANT: Lemarco, Kelly

APPLICANT: Wilson, Adams

APPLICANT: Hertz, Winship

TITLE OF INVENTION: A NOVEL PHARYNGEAL TRANSCRIPTION PROTEIN

TITLE OF INVENTION: HOST CELL FACTOR

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLERR, HOBACH, TEST, ALBERTSON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/393,703

FILING DATE: 24-FEB-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A.

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-57503-2/RAG

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 272299

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 2035 amino acids

```

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-393-703-5

Query Match:
17.8%; Score 67; DB 1; Length 2035;
Best Local Similarity 32.8%; Pred. No. 1e+02;
Matches 22; Conservative 10; Mismatches 27; Indels 8; Gaps 3;

QY      2 PPAPQVDSIQVHSSQPSG-----QAVTVSRQPSLNAYNSLTR--SGLKRTPSLKPDVPP 54
Db      1741 PSTVALLPSTATSLAPSNITVAPOPVVVASPAKIOAAALTEVANGIE-SLGVKPDIPP 1799

QY      55 KPSFAPL 61
Db      1800 PPSKAPM 1806

```

Search completed: September 30, 2003, 16:36:16

Cdb time : 43 secs